

Big data a elektronické zdravotníctvo

Zber a spracovanie veľkých súborov dát možno považovať za jeden z najvýznamnejších technologických trendov súčasnosti. Technológia veľkých dát so sebou prináša pozitívne i negatívne dopady. Tie pozitívne možno vidieť v tom, že nám veľké dáta dávajú do rúk nástroj na to, aby sme dokázali lepšie monitorovať vývoj sveta, predikovať a predchádzať možným problémom. Negatívom sú najmä všetky problémy spojené s možnosťami kontroly správania sa človeka ako jednotlivca. To má za následok potrebu znova sa zamýšľať nad významom a možno aj predefinovať také tradičné hodnoty ako sú sloboda, súkromie, diskriminácia. Jednou z najvýznamnejších oblastí, v ktorej sa veľké dáta s úspechom stále intenzívnejšie využívajú, je oblasť zdravotníctva. Elektronické zdravotníctvo má potenciál zlepšiť zdravotnú starostlivosť. Vývoj systémov a aplikácií už v tejto oblasti značne pokročil a s ohľadom na globalizáciu súčasnej spoločnosti je dôležité, aby sa aj tu postupovalo na základe rešpektovania princípov koordinácie a štandardizácie, tak aby bola umožnená vysoká miera interoperability systémov.

Informácie vždy boli v histórii spoločnosti cenenou komoditou. Kto mal informácie, kto vedel aspoň niečo málo o princípoch, na ktorých svet funguje, mal moc nad ostatnými, pretože dokázal prírodné či spoločenské javy predvídať, alebo nebudaj usmerňovať. To bola úloha žrecov a šamanov v kmeňovej spoločnosti, o tom bolo do značnej miery aj postavenie vzdelanejších vrstiev obyvateľstva v stredoveku. Význam informácií podstatne narástol v priebehu rozvoja industriálnej spoločnosti, predovšetkým od konca 19. storočia, kedy sa informácie začínajú spájať s takými fenoménmi ako hospodárska súťaž, konkurencia, reklama. V priebehu 20. storočia, najmä v jeho druhej polovici, dochádza na vývinovej trajektórii ľudstva k ďalšiemu zásadnému posunu, od industriálnej spoločnosti smerom k spoločnosti informačnej, resp. znalostnej – teda takej, v ktorej sa informácie, ich produkcia a trhová realizácia stávajú výrobnou silou, ekonomickým odvetvím, ktoré hrá stále významnejšiu úlohu v celkovom obraze hospodárstva.

V predchádzajúcich historických etapách išlo o informácie, ktoré dominantne zhromažďovali, spracúvali a rozširovali ľudia prostredníctvom formálnych organizácií, inštitúcií a systémov. A hoci sa v týchto systémoch od cca polovice 20. storočia stále významnejšie pri procesoch spracúvania informácií využívali prostriedky výpočtovej techniky, pri tvorbe informácií stále bol dominantným prvkom človek. Táto situácia sa začína postupne meniť ku koncu 20. storočia, kedy počítače a komunikačné siete pomaly zasahujú aj do procesov zberu a komunikácie dát a postupne prerastá do éry, kedy aj produkcia dát/informácií prestáva byť doménou cieľavedomej ľudskej činnosti.

Pojem *veľké dáta*, či *big data*, sa v súčasnosti v rôznych vedných oblastiach skloňuje čoraz častejšie. S rozvojom technológií a internetu a s nástupom informačnej spoločnosti sa objem dát radikálne zvýšil do takej miery, že nie je možné ich spracovať bežne používanými hardvérovými a softvérovými prostriedkami v reálnom čase. Niektoré súčasné odhady hovoria, že spoločnosť denne vyprodukuje 2,5 trilióna (10^{18}) bytov rôznych typov neštruktúrovaných dát, napríklad posielaním fotografií a videí, realizovaním nákupov cez internet, vyhľadávaním na webe, využívaním aplikácií či GPS navigácií. Tieto dáta zbierajú a zhromažďujú rôzne spoločnosti s cieľom využiť ich vo svojej činnosti, napríklad pri zlepšení marketingových stratégií, urýchlení procesov riadenia, rozhodovania a implementácie riešení, alebo pri zlepšovaní poskytovaných služieb.

Cieľom tohto príspevku je zosumarizovať základné charakteristiky fenoménu veľkých dát a pozrieť sa na ich širšie spoločenské súvislosti. Na základe toho sa zameriame bližšie na problematiku elektronického zdravotníctva (e-health) ako jedného z najaktuálnejších trendov „informatizácie“ spoločnosti, v kontexte zbierania a spracovania veľkých dát. Zmapujeme základné prednosti a spoločenské riziká využívania tejto technológie v oblasti zdravotníctva a spomenieme legislatívne a štandardizačné súvislosti veľkých dát s ohľadom na potrebu medzinárodnej koordinácie ich tvorby a využívania.

Big data

Termín *big data* sa v súčasnej odbornej literatúre nevyvedzuje jednoznačne. Vo vedeckej sfére sa termín prvýkrát objavuje v roku 1997, avšak až v posledných rokoch ide o skutočný fenomén, nakoľko zásadným spôsobom ovplyvňuje spoločnosť. Momentálne sa termín nevzťahuje iba na samotné dáta a ich kvantitu, ale ďalšie definície hovoria o schopnosti spoločnosti využívať informácie novým spôsobom, produkovať užitočné poznatky alebo tovar a služby významnej hodnoty.

Edd Dumbill charakterizuje *big data* ako „dáta, ktorých spracovanie prekračuje kapacitu konvenčných databázových systémov. Dáta sú veľmi veľké, pohybujú sa veľmi rýchlo alebo ich štruktúra sa nehodí do existujúcej databázovej architektúry“. Autor zároveň zdôrazňuje, že ak chceme získať hodnoty z týchto dát, je nevyhnutné zvoliť alternatívny spôsob, ako ich spracovať (Dumbill, 2012).

Big data sa od prvotných pokusov o definovanie veľmi často charakterizujú ako fenomén, pre ktorý sú typické vlastnosti 3V – objem (volume), rýchlosť (velocity) a rôznorodosť (variety). V novších definíciách sa už často objavujú aj ďalšie „V“ charakteristiky, označujúce veracity (vierohodnosť) (Big data, 2015), value (hodnota) či variability (premenlivosť, nekonzistentnosť) (napr. Savic, 2018):

- *objem*: množstvo denne vygenerovaných dát narastá exponenciálne;
- *rýchlosť*: so zvyšujúcim sa objemom dát rastie rýchlosť ich vzniku a potreba ich spracovania, pričom sa vyžadujú analýzy týchto dát v reálnom čase a očakáva sa okamžitý prístup k výsledkom spracovania;
- *rôznorodosť*: generované dáta pochádzajú z rôznych zdrojov a až 80 % dát je neštruktúrovaných;

- **vierohodnosť:** dôležitá je aj pravdivosť spracovaných dát, vzhľadom k tomu, že sú často podkladom na riadenie a rozhodovanie.

Z hľadiska spôsobu ich vzniku, resp. ich zdroja, možno údaje, ktoré sú súčasťou konceptu technológie *big data*, rozdeliť do niekoľkých skupín:

- **dáta z pozorovania** – prístrojovo generované dáta, dáta v reálnom čase, napríklad zo senzorov a snímačov a pod. Hlavným zdrojom dát z pozorovaní je *Internet of Things* (IoT), teda *internet vecí*. Ide o pomerne nový trend v oblasti kontroly a komunikácie objektov rôzneho využitia medzi sebou navzájom, alebo medzi prístrojmi a človekom, a to najmä prostredníctvom technológií bezdrôtového prenosu dát a internetu.
- **interaktívne dáta** – dáta z rôznych sociálnych sietí a sledovania aktivity na internete, ako aj samotný obsah – web logy, videá, fotografie a pod.
- **transakčné dáta** – dáta, ktoré popisujú nejakú udalosť, zmenu ako výsledok transakcie. Vždy obsahujú zápis o čase vzniku prípadne zmeny, numerickú hodnotu a referujú o jednom alebo viacerých objektoch. Typickým príkladom sú napríklad: finančné operácie, logistické operácie alebo zápisy pracovnej aktivity. (Suja, 2014)

Podľa ďalšej kategorizácie by sme mohli *big data* rozdeliť na:

- **štruktúrované** – dáta ktoré sú uložené v presne definovaných a popísaných dátových poliach (typickým príkladom je databáza zákazníkov);
- **neštruktúrované** – tieto nemajú presne definovanú štruktúru (napr. obrázky, videá);
- **pološtruktúrované** – sú prienikom vyššie spomínaných typov (napr. e-mail, ktorý môže mať štruktúrovanú hlavičku ale neštruktúrovaný obsah).

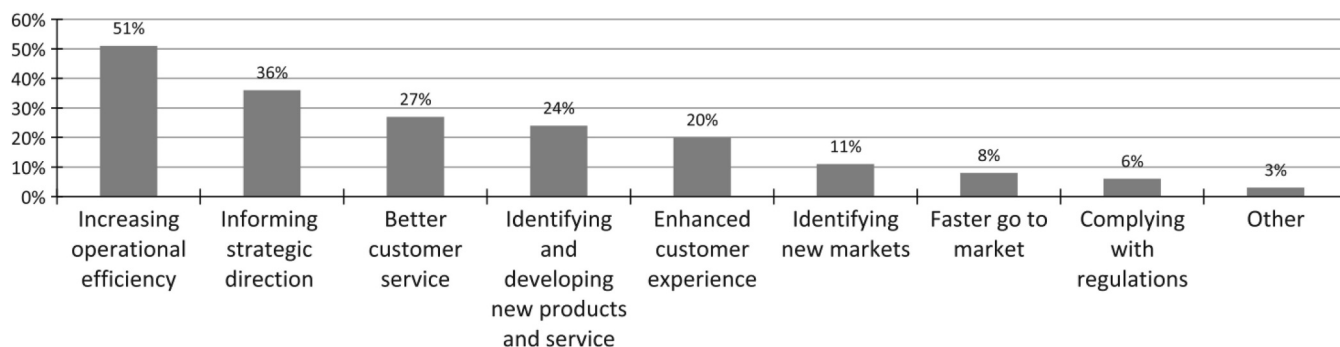
Ako vo svojej štúdii zdôrazňujú Metcalf, Keller a Boyd, „veľkosť“ veľkých dát spočíva v stále sa rozširujúcej schopnosti prepájať separátne dátové sady prostredníctvom algoritmických analýz, čím sa vytvárajú nepredvídateľné vzťahy medzi dátami zozbieranými v rozličných časoch, na rozličných miestach a na rozličné účely (Metcalf et al, 2016). Túto charakteristiku výstižne dopĺňa Rob Kitchin, keď konštatuje, že na rozdiel od tradičných dátových sád, ako sú napríklad hromadné sčítania ľudu, *big data* sú typické tým, že sa generujú kontinuálne, v snahe o zachytenie vyčerpávajúceho a zároveň podrobného obrazu skúmanej reality, a tak aby ich produkcia bola flexibilná a prispôsobiteľná (Kitchin, 2014).

Využitie big data

Z hľadiska využitia môžeme vo sfére veľkých dát rozlíšiť tri základné aplikačné oblasti: marketing a podnikanie, spoločnosť, a vedecký výskum. Každá z týchto oblastí pracuje s určitým typom informácií, ktoré sa vyznačujú vlastným súborom problémov pri snahe o ich využitie a spracovanie.

V oblasti *marketingu a podnikania* je to obrovské kvantum dát, ktoré vzniká po celom svete a tieto dáta reflektujú účel podnikania, aktivity podniku a aktivity zákazníkov. Firmy môžu využívať tieto dáta napríklad na predikovanie vývoja.

Správa z McKinseyovho inštitútu, ktorý sa venuje výskumu v oblasti globálnej ekonomiky, tiež poukazuje na dôležitosť veľkých dát. Obrázok 1 zobrazuje výsledky prieskumu zameraného na identifikáciu oblastí využitia veľkých dát. Viac ako 50 % z 560 opýtaných spoločností si myslí, že im *big data* pomôžu zo zlepšením efektívnosti prevádzky. Na druhom mieste sa s 36 % umiestnila podpora strategického riadenia vďaka veľkým dátam. Určitý význam vidia spoločnosti aj pri zlepšovaní existujúcich (27 %) a pri vývoji nových produktov a služieb (24 %). Do pozornosti sa dostáva aj zlepšenie ako zákazníckych služieb, tak aj zákaznickej skúsenosti. (Chen a Zhang, 2013)



Obr. 1 Príležitosti v oblasti big data (podľa Chen a Zhang, 2013)

Sociálne dáta zbierané a analyzované vo veľkom rozsahu majú obrovský potenciál zmeniť spôsob fungovania spoločnosti a aj kvalitu života. Ich segmentácia na vekové skupiny môže poskytnúť napríklad informácie dôležité pre skvalitňovanie vzdelávacieho systému (v prípade mladších vrstiev populácie), alebo zdravotnej starostlivosti (pri starších ľuďoch).

V oblasti *vedy a výskumu* sú veľké dáta v súčasnosti už známou témou. Mnohé vedné odbory, ako napríklad astronómia, meteorológia, bioinformatika a podobne, sú dnes do istej miery postavené na zbere a spracovaní veľkých súborov dát. Výhodou dát v oblasti výskumu je možnosť ich viacnásobného využitia, napríklad tak, že štandardný experiment sa dá opakovane usku-

točniť na danej množine dát. Dáta je tiež možné zdieľať medzi rôznymi inštitúciami a umožniť tak viacerým výskumným tímom pracovať na identickej téme, alebo objaviť v danej množine dát niečo úplne odlišné (Chen a Zang, 2013).

V oblasti vedeckého výskumu sa v ostatných rokoch rozbehlo niekoľko významných medzinárodných projektov, ktoré sú postavené na využívaní veľkých dát. *International Barcode of life* je dátový systém založený na princípe získavania, ukladania, analýzy a publikovania záznamov DNA. Ide o projekt konzorcia inštitúcií z 25 štátov, ktorých úlohou je analýza genómu DNA za účelom globálnej ochrany ekosystémov. S ich pomocou dokážu profesionálni aj amatérski vedci lepšie mapovať ohrozené druhy, ale aj zdieľať a analyzovať ich DNA informácie. Aplikácie a mobilné technológie, založené na *big data*, ako aj sekvenčné spracovanie DNA ďalšej generácie sa použijú pri tvorbe platformy na skúmanie vzoriek DNA v reálnom čase kdekoľvek vo svete.

Hadoop je jedným z novších príkladov spracovania veľkých dát. Ide o open source framework na spracovanie, ukladanie a analýzu veľkého množstva distribuovaných a neštruktúrovaných dát. Jeho úlohou je pristupovať k dátam dostupným na webe alebo v dátových centrách a rozdeľovať ich do replikovaných častí a tie spracovávať na jednotlivých uzloch. Ďalšie dotazy sa potom spracúvajú paralelne v každom uzle. Výsledky sa agregujú a ukladajú do úložnej vrstvy nazývanej HDFS (Prajapati, 2013). Hlavnou výhodou Hadoopu je, že umožňuje analyzovať úplné dátové súbory vrátane neštruktúrovaných alebo čiastočne štruktúrovaných dát, a to efektívne z hľadiska času i nákladov.

Veľký hadrónový urýchľovač (LHC) je urýchľovač častíc, ktorý bol skonštruovaný vo švajčiarskom CERN-e. Výsledky experimentov, ktoré sa tu uskutočňujú, prispievajú k lepšiemu chápaniu fyzikálnych interakcií. Jedným z cieľov LHC je zodpovedanie fundamentálnych otázok týkajúcich sa elektromagnetickej a jadrovej interakcie, gravitácie a symetrie. Dôležitú úlohu pri spracovaní veľkých objemov dát hrá vedecká spolupráca. Pri experimentoch na urýchľovači v CERN-e vzniká obrovské množstvo vedeckých dát, ktoré ani najrýchlejšie prístroje nie sú schopné dostatočne rýchlo spracovávať. Aj po spracovaní tieto údaje potrebujú dodatočnú analýzu a najmä interpretáciu zo strany vedcov. Rovnako ako pri analýze údajov z LHC, aj pri analýze údajov v astronómii vznikajú kolaborácie vedcov z celého sveta, najmä s cieľom zrýchliť proces analýzy.

Zobierané dátové sady môžu mať veľký význam aj v iných oblastiach vedy a výskumu, ako sú napríklad spoločenskovedné alebo humanitné disciplíny. Hoci v minulosti to bolo len ťažko predstaviteľné, dnes tieto sady odborníci využívajú i pre potreby kvantitatívnych výskumov vo sfére politológie. Môže ísť napríklad o analýzu príspevkov na webe počas volebného obdobia, kedy vedci využijú matematické algoritmy na zistenie preferencií voličov. Povaha týchto výskumov v rámci sociálnych vied je veľmi špecifická. Potrebujeme na ne technológie na využitie *big data*, ale aj odborníkov schopných realizovať kvalitatívny výskum v danej oblasti, ktorí po analyzovaní dát prinášajú nové zistenia a závery.

Zdravotníctvo a big data

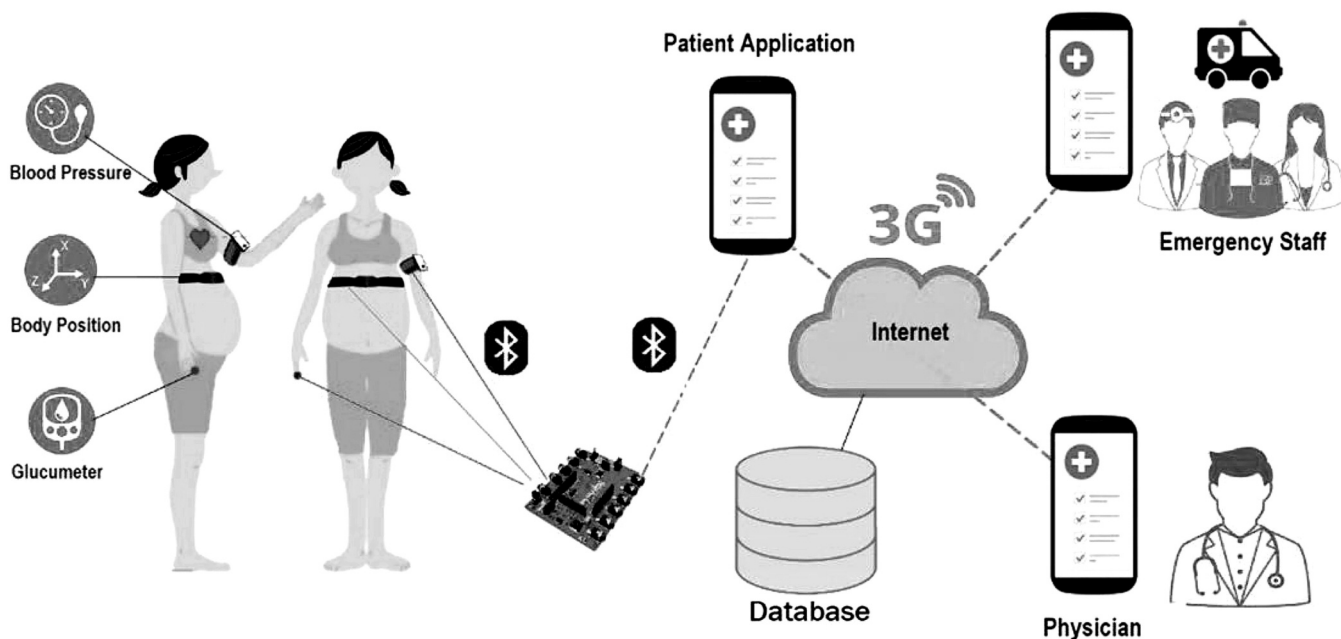
Jednou z oblastí, kde veľké dáta nachádzajú svoje uplatnenie, je sektor zdravotníctva. V tejto oblasti predstavujú *big data* riešenie na zlepšenie efektivity pri spracovaní a využívaní zdravotníckych informácií. Elektronické zdravotníctvo je fenoménom, ktorý sa v ostatných rokoch dostal do popredia záujmu médií, politik i širokej verejnosti, ako jedna z praktických aplikácií informačných technológií do bežného života. Už v roku 2005 vo svojom prehľadovom článku pre *Journal of Medical Internet Research* Pagliari et al. konštatovali, že počet publikácií, ktoré sa venujú problematike elektronického zdravotníctva, začal prudko stúpať od roku 2000 (Pagliari, 2005). E-health pritom zahŕňa také oblasti ako nástroje na podporu rozhodovania, elektronické zdravotné záznamy, telemedicína a aplikácie starostlivosti na diaľku, elektronické predpisovanie liekov a objednávanie pacientov či mobilné zdravotné zariadenia, postavené okrem iného na masovom rozšírení mobilných komunikačných zariadení v ostatných desaťročiach.

Zavedenie a využívanie digitálnych technológií do oblasti zdravotníctva môže pomôcť skvalitniť život pacienta a lepšie chrániť jeho zdravie v mnohých ohľadoch. Tým, že sa navzájom prepoja jednotlivé databázy údajov a zlepši sa vzájomná informovanosť lekárov o stave pacienta, môžu rýchlejšie identifikovať jeho diagnózu, individuálne prispôbiť zdravotnú starostlivosť, nastaviť mu účinnejšiu liečbu, a tým predchádzať vzniku ďalších ochorení. Lekári budú mať k dispozícii kompletnú elektronickú lekársku správu pacienta. Budú vidieť akých lekárov navštevuje, ktoré lieky užíva, v akom počte a v akom intervale, kedy bol naposledy na preventívnom vyšetrení a mnohé iné skutočnosti, na ktoré sa v prípade lekárskej evidencie v papierovej forme často zabúda alebo sa zámerne zatajújú (Viola, 2016).

Primárnym cieľom využitia *big data* v zdravotníctve je predpovedanie epidémií, liečenie chorôb, zlepšenie kvality života, redukcia chýb pri určení diagnózy a hlavne vyvarovanie sa zbytočným úmrtiam spozorovaním varovných príznakov problémov v počiatočnom štádiu (Marr, 2015). V súčasnosti existuje množstvo mobilných aplikácií, ktoré do istej miery mapujú časť informácií o zdraví človeka. V budúcnosti aj tieto údaje môžu byť dôležitou súčasťou zdravotníckych informácií o pacientovi. Primárnym účelom mobilných aplikácií je to, aby pomohli svojim používateľom žiť zdravšie a zlepšiť svoj životný štýl. Informácie z týchto aplikácií, ktoré ľudia využívajú, či už na zaznamenávanie počtu krokov, dodržiavanie pitného režimu alebo naplánovanie jedálneho lístka, bude v budúcnosti možné zdieľať so svojim lekárom a lekárnikom (Marr, 2015).

K údajom v elektronickej forme budú mať pravdepodobne prístup aj osoby, ktoré aktívne spolupracujú s odvetvím zdravotníctva. Na jednej strane budú tieto údaje cenným zdrojom informácií, ale na druhej strane predstavujú riziko. Organizácie budú môcť vidieť kompletné zdravotné informácie pacienta, a ak zistia, že jednotlivec má minimálny denný výdaj energie, nevenuje sa žiadnej aktivite, často konzumuje jedlá z kategórie rýchleho občerstvenia a pije vysoko kalorické nápoje, môže napríklad poisťovňa odmietnuť jeho žiadosť o uzatvorenie poistenia alebo mu môže zvýšiť výšku poistenia.

Potvrďuje sa tak to, čo už pred pár rokmi v súvislosti s veľkými dátami vo všeobecnosti (nielen v zdravotníctve) konštatovali Crawford a Schultz – veľké dáta posúvajú naše vnímanie etiky vo výskume od tradičných kategórií „ujmy“, ako sú fyzická bolesť či psychický stres, k iným kategóriám ako trvalý dohľad, individuálna alebo skupinová diskriminácia a prediktívne ujmy na súkromí, kde sa zásahy do privátnej sféry uskutočňujú na základe odvodzovania, a nie priameho zberu osobných dát (Crawford, 2013).



Obr. 2 Schéma fungovania zdravotníckej aplikácie založenej na zbere dát (E-Health, 2017)

Uplatnením metód spracovávaní *big data* v zdravotníctve sa vyrieši problém s obrovským množstvom denne kumulovaných dát, ktoré lekárska prax zhromažďuje, bez následného efektívneho využitia týchto cenných informácií. Snahou je využiť potenciál informácií, ktoré často zostávajú uložené len v archíve lekára v ordinácii, v nemocnici, na klinike či zariadeniach pacientov. Ako príklad možno uviesť *Smart m-Health Application for Pregnancy Care Using Body Area Networks*, ktorá vychádza zo skúsenosti, že jednými z najčastejších problémov žien v tehotenstve sú vysoký tlak a diabetes. Aplikácia za pomoci mobilných zariadení monitoruje aktuálne hodnoty vybraných fyziologických premenných u pacientky, prenáša ich do databázy, kde sa spracúvajú a vyhodnocujú a v prípade potreby sa k nim dostane lekár alebo rýchla zdravotnícka pomoc.

Výhody budú spočívať taktiež v úspore finančných nákladov, keď sa na základe správne nastavenej účinnej a efektívnej liečby môže skrátiť dĺžka hospitalizácie. Zníženie finančných nákladov môže predstavovať aj odstránenie duplicitných vyšetrení pacienta alebo úsporu financií potrebných na administratívne záležitosti.

Právna ochrana a štandardizácia dát v zdravotníckych systémoch

Rozvoj elektronických služieb a využívanie *big data* v oblasti zdravotníctva priniesli už od počiatku otázky právnej a technologickej ochrany citlivých dát o zdravotnom stave pacientov. Elektronizácia zdravotníctva si v mnohých krajinách vyžiadala vznik novej legislatívy v oblastiach ochrany osobných zdravotných informácií, ktorá reguluje primárne a sekundárne využívanie zdravotných informácií. Medzi hlavné nástroje ochrany dát v systéme elektronického zdravotníctva zaraďujeme zákony, bezpečnostné normy a predpisy (Shoniregun, 2010). Všetky tieto tri nástroje sa navzájom dopĺňajú pri ochrane týchto dát. Z hľadiska ochrany dát je nielen pre zdravotnícke dáta najdôležitejšie spomenúť Smernicu 1995/46/ES, ktorej obsah súvisí s ochranou osobných údajov a súkromia. Smernica:

- stanovuje procesy nevyhnutné pri spracovaní osobných údajov;
- zakazuje odosielanie osobných údajov z krajín EÚ do takých krajín, ktoré neposkytujú primeranú ochranu osobných údajov;
- vyzýva členské štáty na prijatie primeraných opatrení pre zabezpečenie vykonávania tejto smernice.

Otázku bezpečnosti dát v systéme elektronického zdravotníctva na Slovensku pokrývajú primárne dva zákony – zákon č. 122/2013 o ochrane osobných údajov a zákon č. 153/2013 o národnom zdravotníckom informačnom systéme. Zákon o národnom zdravotníckom informačnom systéme obsahuje aj paragraf definujúci údaje, ktoré je možné využiť na zdravotnícku štatistiku, ktorá vlastne predstavuje využitie *big data* z údajov slovenských pacientov pre potreby štátnej štatistiky. Tieto údaje budú anonymizované a poskytnúť ich tretím stranám bude možné len na základe žiadosti (Zákon č. 153/2013 Z. z.).

Prevažná väčšina krajín EU v súčasnosti má, alebo v blízkej budúcnosti plánuje implementovať systém elektronického zdravotníctva. S tým súvisí potreba ich potenciálneho vzájomného prepojenia na základe využívania rovnakých štandardov, ktoré zabezpečujú vzájomnú *interoperabilitu* systémov. Interoperabilita, a to nielen v oblasti zdravotníckych systémov, má za cieľ:

- zvýšiť flexibilitu systémov, na základe ktorej vznikne možnosť ich kombinovania;
- zvýšiť efektívnosť nákladov;
- vytvorí vzájomné integrované systémy, ktoré sú ľahšie využiteľné medzinárodne;
- uľahčí vytváranie nových funkcií, založených na existujúcich funkciách systémov (European Commission, 2015).

S cieľom zvýšenia interoperability systémov bol v rámci Európskej únie vyvinutý Európsky rámec interoperability (European Interoperability Framework). Aktuálne sa využíva jeho druhá verzia, ktorú Európska komisia prijala 16. decembra 2010. Cieľom rámca je presadiť a podporiť poskytovanie európskych verejných služieb prostredníctvom podpory cezhraničnej a medzirezortnej interoperability, presadiť poskytovanie verejných služieb podnikom a občanom prostredníctvom orgánov verejnej správy a doplniť a prevziať jednotlivé národné rámce na európskej úrovni (European Commission, 2015). Pri prezentácii druhej verzie rámca bol predstavený aj nový rámcový model, ktorý pozostáva zo šiestich úrovní:

- Legislatíva a regulácia – upravuje interoperabilitu na úrovni práva a regulačných zásad medzi krajinami, ale aj v rámci krajiny alebo regiónu;
- Politika – predstavuje zmluvy a dohody medzi organizáciami o vzájomnej spolupráci;
- Proces starostlivosti – po uzatvorení dohôd sa vykonáva analýza procesov v rámci organizácie a ich zladenie;
- Informácie – časť predstavujúca popis dátového modelu, dátových prvkov a vzájomné prepojenie prvkov;
- Aplikácie – predstavuje proces importu a exportu informácií medzi systémami, ale aj prípravu grafického rozhrania pre využívanie systémov;
- IT infraštruktúra – súpis sieťových protokolov, štandardov, uchovávaní dát, databázových systémov a podobne.

Len vďaka interoperabilným systémom elektronického zdravotníctva bude možná vzájomná výmena *big data* na medzinárodnej úrovni, vďaka čomu bude možné analyzovať dáta aj na globálnej úrovni. Výsledky analýz umožnia predpovedať a vytvárať analytické štúdie napríklad o zdravotnom stave ľudí v Európskej únii, či zistiť najčastejšiu príčinu úmrtia pacientov v Európskej únii a na základe toho vypracovať akčné plány na zníženie rizikových faktorov a podobne.



Obr. 3 Príklad aplikácie modelu interoperability medzi dvoma organizáciami (European Commission, 2015)

Okrem štandardizácie a interoperability systémov elektronického zdravotníctva je potrebné aj pevné stanovenie oblastí, kde sa budú *big data* využívať. V Európskej únii s týmto cieľom bola vypracovaná štúdia o využívaní veľkých dát vo verejnom zdravotníctve, telemedicíne a zdravotnej starostlivosti. Tá bola v decembri 2016 predstavená na pôde Európskej komisie a jej cieľom bolo nájsť reálne využiteľné metódy spracovania *big data* v oblasti zdravotných údajov a odporučiť metódy vhodné na realizáciu v rámci EÚ.

Štúdia skúmala 20 rôznych možností využitia veľkých dát, pričom každý skúmaný systém alebo spôsob využitia dát bol posudzovaný z hľadiska pridanej hodnoty pre pacienta, poskytovateľa zdravotnej starostlivosti, politiky, výskumu a vyhodnotená bola aj celková pridaná hodnota systému. Okrem toho štúdia bola významná aj z dôvodu vytvorenia desiatich odporúčaní pre politiku využívania zdravotníckych *big data* v Európskej únii. Tieto odporúčania sa týkajú nasledovných tematických okruhov:

- zvyšovanie informovanosti: vytvoriť a vykonávať komunikačnú stratégiu na zvýšenie povedomia o pridanej hodnote spracovania veľkých dát v oblasti zdravotníctva a podporovať pozitívne verejnú mienku na spracovanie dát v oblasti zdravotníctva;
- vzdelávanie a odborná príprava: posilniť ľudský kapitál vzhľadom na nárast potreby pracovných síl, ktoré sa môžu využiť na spracovanie *big data* v zdravotníctve;
- dátové zdroje: rozšírenie existujúcich a preskúmanie nových zdrojov dát v oblasti zdravotníctva a zabezpečenie ich kvality a bezpečnosti;
- otvorené dáta a zdieľanie dát: podporovať otvorené využívanie a zdieľanie *big data*, bez ohrozenia práv pacientov na súkromie;
- aplikácie a použitie: cielená aplikácia analýz *big data* v oblasti zdravotníctva na základe potreby a záujmov zainteresovaných strán vrátane pacientov;

- analýzy dát: stanovenie potenciálu analýzy veľkých dát, zlepšenie analytických metód a uľahčenie využívania nových a inovatívnych analytických metód;
- prístup k dátam a ich využitie: zaviesť mechanizmus riadenia s cieľom vytvoriť bezpečný a spravodlivý prístup a využívanie *big data* pre výskumy v oblasti zdravotníctva;
- normy: vypracovať štandardy pre spracovanie veľkých dát v oblasti zdravotníctva, zlepšiť a zjednodušiť ich uplatňovanie a zlepšiť interoperabilitu;
- financovanie a finančné zdroje: zabezpečiť investície riadené Európskou komisiou, ktoré budú nákladovo efektívne a udržateľné;
- právne aspekty a predpisy o ochrane súkromia: vyjasniť a zjednotiť právo a zásady regulácie spracovania *big data* v zdravotníctve (European Commission, 2016).

Problémy vo využívaní big data

S rastúcim množstvom ukladaných dát logicky rastú nároky aj na ich spracovanie. V súčasnom svete už tieto nároky nie sú ani tak o kvantite, ale skôr o kvalite spracovania a o tom, čo z daných dát dokážeme vyťažiť. Lídrmi v spracovaní *big data* sú krajiny ako USA a ich Národný bezpečnostný úrad, v Ázii Južná Kórea a Japonsko (Suja, 2014). V Európe sa problematikou veľkých dát zaoberá *Európsky dátový portál* (2017). Jeho úlohou je zbieranie metadát z verejného informačného sektora a verejných portálov dát vo všetkých krajinách Európskej únie. Ďalšími projektmi, ktoré sa zaoberajú problematikou *big data* v rámci EÚ, sú Big Data Europe (2017) a Digital Single Market (2016).

Pri spracovávaní dát je predovšetkým potrebné mať jasne definové, čo chceme zistiť. V súčasnosti prostredníctvom *big data* dokážu úspešne sociálne aplikácie odhaliť potenciálnych teroristov, odhaliť finančnú situáciu používateľa, ako aj prispôbiť reklamu potrebám používateľa. Napríklad PRISM program sa snaží zistiť cez *big data*, kto sa pokúšal kontaktovať známych teroristov cez SMS alebo telefón. Systémy sa tiež učia rozoznávať, čo sa nachádza na obrázkoch (Ni, 2016).

Na spracovanie veľkých súborov dát pritom nie je potrebné zaobštarávať si nové, drahé zariadenia alebo sa učiť nové postupy. Dokážeme ich analyzovať aj tradičnými nástrojmi. Dôležité je, aby analýzu dát vykonávala osoba, ktorá je na to školená. Väčšinou touto osobou je *informačný analytik*. Informační analytici dokážu predpovedať problémy, ktoré môžu nastať pri určitých procesoch. Dokážu však aj navrhnúť opatrenia na ich zmiernenie a umožňujú efektívnu implementáciu procesov. Takúto činnosť informačných analytikov nazýva Stojanovič (2015) kreatívnou analýzou, keďže informační analytici dokážu zvýšiť tvorivý potenciál pre zlepšenie procesov vo firme. K pochopeniu takýchto procesov potrebujú využívať tvorivé myslenie. Kreatívni analytici majú vo firmách ale aj verejnom sektore veľké využitie. Ako pri spracovaní veľkého objemu dát, tak pri analýze toho, ako zlepšiť obchodné modely, či toho, ako vytvoriť segmentáciu zákazníkov. Informačný analytik môže vykonávať celý rad sofistikovaných úloh, ktoré priamo súvisia s informačnými potrebami používateľa a správou veľkých súborov dát.

V budúcnosti je ale pravdepodobné, že úloha informačného analytika v týchto procesoch sa bude postupne zmenšovať. Už teraz totiž vznikajú rôzne vízie o budúcnosti spracovania *big data*. Ni (2016) píše, že ďalším krokom pri vývoji a spracovaní dát budú systémy *analytics-as-a-service*. Budú to systémy, do ktorých budú vstupovať surové dáta, a následne z nich vychádzajú spracované zanalyzované výsledky.

Výhod, ktoré vyplývajú zo zavedenia *big data* do oblasti zdravotníctva, je veľa. Ako však už bolo viackrát zdôraznené, treba zvážiť aj negatívne aspekty v kontexte ochrany súkromia a bezpečnosti osobných údajov. Pokiaľ nebudú vyriešené základné otázky bezpečnosti, nemožno predpokladať úspešné využívanie týchto systémov.

Okrem štúdie, ktorú si nechala vypracovať Európska komisia o možnostiach využitia *big data* v zdravotníctve, podobnej problematike sa venovala aj Svetová zdravotnícka organizácia (WHO), ktorá sa zamerala na zavádzanie e-health systémov v Európe. Zostavila štúdiu na základe vyhodnotenia dotazníkov, ktoré rozposlala 53 členským štátom WHO pre Európsky región. Vo výskume sa WHO zamerala aj na problematiku implementovania *big data* v zdravotníctve. Na základe 44 odpovedí členských krajín vyhodnotila najčastejšie problémy pri implementovaní tejto technológie v zdravotníctve. Najväčšou prekážkou pri využívaní veľkých dát boli obavy o súkromie a bezpečnosť, nasledovala nedostatočná integrácia, prekážky pri zdieľaní informácií, presadzovanie štandardov, nedostatočné kapacity pri využívaní *big data* a potreba podpory nových analytických metód (World Health Organisation, 2016).

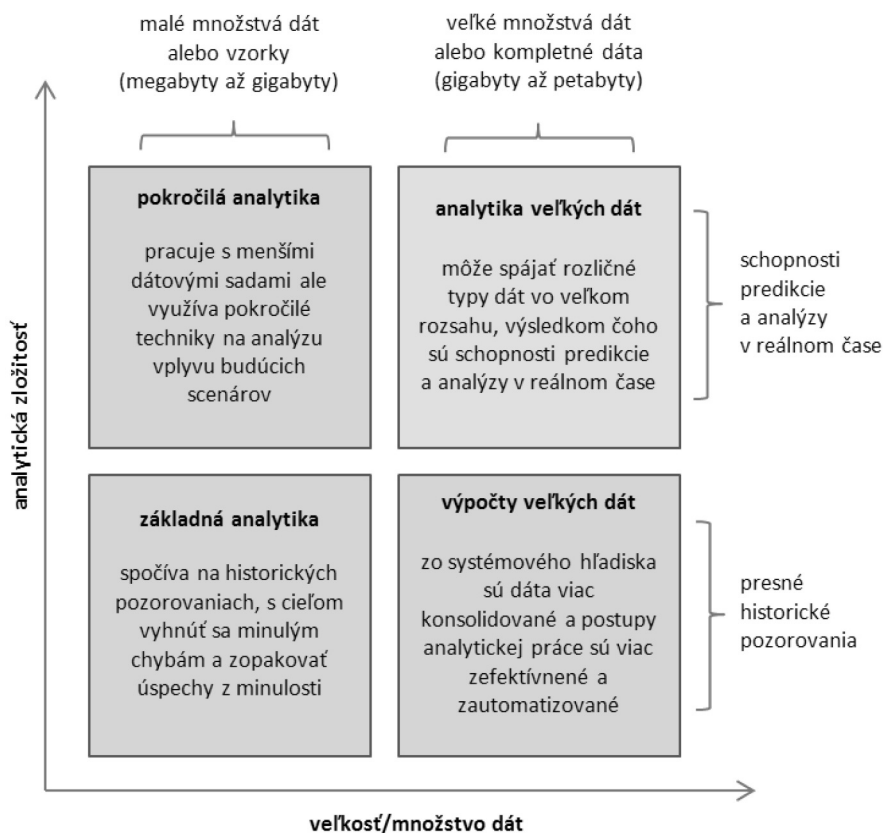
V tejto súvislosti je potrebné vytvoriť legislatívne a právne opatrenia na národnej aj medzinárodnej úrovni. Jednoznačne identifikovať, kto bude mať k informáciám prístup, aký bude ich účel využívania a zdôrazniť bezpečný spôsob zaobchádzania s informáciami, ktorý bude rešpektovať súkromie jednotlivca a ochranu jeho osobnosti.

Ďalším dôležitým predpokladom pre správne využívanie *big data* v oblasti medicíny je špecializácia odborníkov, ktorí práve v tejto oblasti potrebujú spĺňať tri aspekty spôsobilosti:

- odborná spôsobilosť – dátoví špecialisti sa potrebujú vedieť orientovať v oblasti medicíny;
- matematická spôsobilosť – schopnosť pracovať na teoretickej úrovni a vytvárať dátové modely, na základe ktorých sú dáta neskôr spracovávané;
- počítačová spôsobilosť – potreba navrhovať a vytvárať softvérové riešenia, ktoré budú vyhodnocovať dostupné dáta.

Z hľadiska využitia dát je taktiež potrebný posun do oblasti analytiky veľkých dát (obrázok 4), v rámci ktorej je možné spracovať obrovské množstvo dát a zároveň zachovať schopnosti predikcie v reálnom čase. V súčasnosti ešte stále využívanie medicínskych dát často funguje na úrovni základnej analytiky, ktorá sa zameriava na minulé pozorovania, učí sa z chýb a snaží sa duplikovať úspechy (Roski, Bo-Linn a Andrews, 2014).

Potreba odstránenia týchto bariér tvorí kľúčový prvok pri rozšírení využívania *big data* v zdravotníctve, ktoré môžu mať výrazne pozitívny vplyv na zdravie pacientov. Už v dnešnej dobe existuje viacero projektov, ktoré na základe zdravotníckych dát predpovedajú stav pacientov do budúcnosti a tým môžu zabezpečiť dlhší a kvalitnejší život s možnosťou eliminácie niektorých zdravotných problémov. V konečnom dôsledku je dôležité vytvoriť u občanov dôveru k systému, zdôrazniť dostatočné zabezpečenie infraštruktúry a vyzdvihnúť pozitíva, ktoré plynú z implementácie veľkých dát v zdravotníctve.



Obr. 4 Prístupy dátovej analytiky, podľa veľkosti dát a zložitosti analýzy (podľa Roski et al., 2014)

Záver

Zbieranie a spracovanie veľkých súborov dát možno považovať za jeden z najsignifikantnejších prejavov štvrtej priemyselnej revolúcie, ktorá v súčasnosti hýbe globálnou civilizáciou. Podstatou tejto revolúcie podľa viacerých teoretikov je to, že do popredia sa dostávajú technológie, ktoré majú schopnosť prepájať fyzické, digitálne a biologické svety. Technológia veľkých dát, ako každá technológia, so sebou prináša pozitívne i negatívne dopady. Tie pozitívne možno vidieť v tom, že nám veľké dáta dávajú do rúk nástroj na to, aby sme dokázali lepšie ako kedykoľvek predtým monitorovať vývoj sveta, v makro- i mikromeradle, pochopiť jeho smerovanie, predikovať a predchádzať možným problémom. Negatívom sú najmä všetky aspekty spojené s možnosťami kontroly správania sa človeka ako jednotlivca. To má za následok potrebu znova sa zamýšľať nad významom a možno aj predefinovať také tradičné hodnoty ako sú sloboda, súkromie, diskriminácia.

Jednou z najvýznamnejších oblastí, v ktorej sa veľké dáta s úspechom stále intenzívnejšie využívajú, je oblasť zdravotníctva. Elektronické zdravotníctvo má potenciál zlepšiť zdravotnú starostlivosť, keďže okamžitý prístup k informáciám o aktuálnom stave pacienta i celkovom vývoji jeho zdravotného stavu možno považovať za jeden z hlavných predpokladov úspešnej liečby. Chawla a Davis (2013) sú vo svojich úvahách o paradigmatických zmenách v zdravotnej starostlivosti ešte radikálnejší a prezentujú model, v ktorom sa zdravotná starostlivosť presúva od tradičného zamerania na chorobu (zdravotný problém – testy – diagnóza – liečenie) smerom k zameraniu na pacienta (individuálny prístup – aktívna účasť pacienta – aktívna prevencia a manažment choroby). Vývoj systémov a aplikácií už v tejto oblasti značne pokročil a s ohľadom na globalizáciu súčasnej spoločnosti – či už berieme do úvahy realitu Európskej únie alebo celého sveta – je dôležité, aby sa aj tu postupovalo na základe rešpektovania princípov koordinácie a štandardizácie, tak aby bola umožnená vysoká miera interoperability systémov.

Príspevok bol spracovaný v rámci riešenia výskumnej úlohy APVV 0508-15 HIBER – Human Information Behaviour in the Electronic Environment/Informačné správanie človeka v digitálnom prostredí.

ZOZNAM BIBLIOGRAFICKÝCH ODKAZOV

Big Data aneb Když běžné databázi dochází dech. In: *IT Systems* [online]. Brno: SystemOnline, 2015, č. 11 [cit. 2018-02-10]. Dostupné: <https://www.systemonline.cz/business-intelligence/big-data-aneb-kdyz-bezne-databazi-dochazi-dech.htm>

- Big Data Europe, 2017 In: *Big Data Europe, Empowering Communities with Data Technologies* [online]. [cit. 2018-02-10]. Dostupné na: <https://www.big-data-europe.eu/>
- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century. In: *European Commission. Digital Single Market* [online]. Brussels: European Commission, 2012 [cit. 2018-02-10]. Dostupné na: <https://ec.europa.eu/digital-single-market/en/news/ehealth-action-plan-2012-2020-innovative-healthcare-21st-century>
- DUMBILL, Edd, 2012. What is big data? An introduction to the big data landscape. In: *O'REILLY* [online]. Sebastopol: O'REILLY, 2012 [cit. 2018-02-10]. Dostupné na: <https://www.oreilly.com/ideas/what-is-big-data>
- E-Health application developed with MySignals first winner in health competition ISHIC 2017. Cooking Hacks. [cit. 2018-02-10]. Dostupné na: https://www.cooking-hacks.com/blog/e-health-application-developed-with-mysignals-first-winner-in-health-competition-ishic-2017/#img_0
- Elektronické zdravotníctvo – eHealth. In: *Národné centrum zdravotníckych informácií* [online]. Bratislava: Národné centrum zdravotníckych informácií, ©2011 [cit. 2018-02-10]. Dostupné na: <http://www.nczisk.sk/eHealth/Pages/default.aspx>
- EUROPEAN COMMISSION, 2015. *EHealth Network: Refined eHealth European Interoperability Framework*. Brussels. [cit. 2018-02-10]. Dostupné na: http://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20151123_co03_en.pdf
- EUROPEAN COMMISSION, 2016. *Study on Big Data in Public Health, Telemedicine and Healthcare: Final Report*. Brussels. [cit. 2018-02-10]. Dostupné na: https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata_report_en.pdf
- CHAWLA Nitesh V. a Darcey A. DAVIS, 2013. *Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework* [online]. *Journal of General Internal Medicine* 28(Suppl 3): 660. 2013. [cit. 2018-02-10]. Dostupné na: <https://link.springer.com/article/10.1007/s11606-013-2455-8/fulltext.html>
- CHEN C.L. Philip a Chun-Yang ZHANG, 2013. *Data-intensive applications, challenges, techniques and technologies: A survey on Big Data* [online]. [cit. 2018-02-10]. Dostupné na: http://www.crisismanagement.com.cn/templates/blue/down_list/llzt_dsj/Data-intensive%20applications,%20challenges,%20techniques%20and%20technologies%20A%20survey%20on%20Big%20Data.pdf
- KITCHIN R., 2014. Big Data, new epistemologies and paradigm shifts. In *Big Data & Society* 1(1): 2053951714528481.
- MARR, Bernard, 2015. How Big Data Is Changing Healthcare. In: *Forbes* [online]. New York: Forbes, 2015 [cit. 2018-02-10]. Dostupné na: <https://www.forbes.com/sites/bernardmarr/2015/04/21/how-big-data-is-changing-healthcare/#464b5cd82873>
- METCALF, Jacob, Emily F. KELLER, and Danah BOYD, 2016. Perspectives on Big Data, Ethics, and Society. In: *Council for Big Data, Ethics, and Society*. [cit. 2018-02-10]. Dostupné na: <https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/>.
- NI, Lionel M., Haoyu TAN a Jiang XIAO, 2016. Rethinking big data in a networked world. In: *Frontiers of Computer Science* [online]. 2016, roč. 10, č. 6, s. 965–967. ISSN 2095-2228, 2095-2236. [cit. 2018-02-10]. Dostupné na: [doi:10.1007/s11704-016-6902-7](https://doi.org/10.1007/s11704-016-6902-7)
- PAGLIARI C., SLOAN D., GREGOR P., SULLIVAN F., DETMER D., KAHAN JP., OORTWIJN W., MACGILLIVRAY S., 2005. What Is eHealth (4): A Scoping Exercise to Map the Field. *J Med Internet Res* 2005;7(1):e9. [cit. 2018-02-10]. Dostupné na: <http://www.jmir.org/2005/1/e9>.
- ROSKI Joachim, BO-LINN George W. a Timothy A. ANDREWS, 2014. *Creating Value In Health Care Through Big Data: Opportunities And Policy Implications* [online]. *Health Affairs* 33, NO. 7: 1115–1122. 2014 [cit. 2018-02-10]. Dostupné na: [10.1377/hlthaff.2014.0147](https://doi.org/10.1377/hlthaff.2014.0147)
- SAVIC, D., 2017. Impact of disruptive Technologies on grey literature management. *IT Lib*, 4/2017, pp. 42-45.
- SHONIREGUN a kol, 2010. *Electronic Healthcare Information Security* [online]. Online-Ausg. Boston, MA: Springer Science Business Media, [cit. 2018-02-10]. ISBN 978-038-7849-195. Dostupné na: http://link.springer.com/chapter/10.1007/978-0-387-84919-5_3
- STOJANOVIĆ, Nenad, 2015. From Big data as "new oil" to Big data as "food for creativity". In: *Digital Single Market* [online]. [cit. 2018-02-10]. Dostupné na: <https://ec.europa.eu/digital-single-market/en/blog/big-data-new-oil-big-data-food-creativity>
- SUJA, Robert, 2014. Big Data. In: *Infostat.sk*. [online] [cit. 2018-02-10]. Dostupné na: http://www.infostat.sk/web2015/sk/_publikacie/Big_Data.pdf
- VIOLA, Roberto, 2016. Digital Solutions for Better Health. In: *European Commission. Digital Single Market* [online]. Brussels: European Commission, 2016 [cit. 2018-02-10]. Dostupné na: <https://ec.europa.eu/digital-single-market/en/blog/digital-solutions-better-health>
- WILDER-JAMES, Edd, 2014. Defining Big Data. In: *Forbes* [online]. New York: Forbes, 2014 [cit. 2018-02-10]. Dostupné na: <https://www.forbes.com/sites/edddumbill/2014/05/07/defining-big-data/#be1ea0956671>
- WORLD HEALTH ORGANISATION, 2016. *From innovation to implementation: eHealth in the WHO European Region (2016)* [online]. 1. Copenhagen [cit. 2018-02-10]. ISBN 978 92 890 5137 8. Dostupné na: http://www.euro.who.int/__data/assets/pdf_file/0012/302331/From-Innovation-to-Implementation-eHealth-Report-EU.pdf

Zákon č. 122/2013 Z. z. Zákon o ochrane osobných údajov a o zmene a doplnení niektorých zákonov.

Zákon č. 153/2013 Z. z. Zákon o národnom zdravotníckom informačnom systéme a o zmene a doplnení niektorých zákonov.

prof. PhDr. Jaroslav Šušol, PhD.

jaroslav.susol@uniba.sk

Mgr. Lukáš Borčín

borcin5@uniba.sk

Mgr. Andrej Garaj

garaj14@uniba.sk

(Katedra knižničnej a informačnej vedy Filozofickej fakulty Univerzity Komenského v Bratislave)

Článok je recenzovaný