Posudek doktorské disertační práce Ing. Richarda Nováka předložené v oboru Aplikovaná informatika na Fakultě informatiky a statistiky Vysoké školy ekonomické v Praze

Oponent: doc. RNDr. Tomáš Pitner, Ph.D.

Aktuálnost tématu

Především bych rád zdůraznil, že téma Big Data and Ethics je vysoce aktuální a rigorózní přístup k němu je nesmírně důležitý. Proto vítám iniciativu studenta, který má potenciál vytěžit ze svých vědeckých i praktických manažerských zkušeností a systematizovat koncepty v předmětné oblasti, jakož i současně odpovědět na vybrané konkrétní otázky, jež považuji za podstatné – jakkoli je jejich relevance vždy subjektivní. Zejména aktuální je zvolený multidisciplinární přístup, který v době rychlého nástupu umělé inteligence do výzkumu i byznysu bude muset být zohledněn v prakticky každém aplikovaném výzkumu zahrnujícím velká data.

Srovnání s mezinárodním výzkumem

Autor věnoval velkou část své výzkumné činnosti právě srovnávání s výsledky ostatních autorů a state-of-the-art. To se týká zejména části teoretických konceptů etiky v oblasti IT a dat. Díky tomu je teoretická báze v práci rozsáhle a fundovaně popsána. Použití statistických metod odráží široký rozhled a schopnost použít potřebné metody kvantitativního výzkumu.

Vědecké metody

Autor využívá kombinaci paradigmat několika vědních disciplín a tyto správně uvádí v definici zkoumané oblasti – empirické metody + logické metody. Z hlediska praktických postupů pak statistickou datovou analýzu dat získaných od uživatelů sociologickou technikou dotazníků.

Záběr disertačního výzkumu autor uvádí v 1.3, zde se jedná hlavně o výstupy popisného typu charakterizované slovy „vyjasnění“, „uspořádání“, „popis“, „diskuse“, „podrobný seznam“, „průzkum“ a také návrhového typu, kde za zásadní vzniklý artefakt považuji uspořádání logického rámce „Regulatory Framework of Big Data“ v kap. 5 s identifikací podstatných legislativních architekturních a uživatelských aspektů využití velkých dat. Zásadnější vlastní závěry staví na formulacích a následném ověřování hypotéz v kap. 6, které pro přehlednost mohly být stanoveny již v kap. 2 a nikoli až 6.3, ale to je jen drobnost.

Ověření platnosti poznatků

V rámci práce vyvinuté koncepty byly motivovány případovými studiemi – „use cases“ v kap. 4. Use cases pokrývaly velkou část domén, kde jsou velká data v současnosti generována – zejména sektor telekomunikací, data o pohybu uživatelů vč. geografických. Na základě nich, a především převzatých teoretických poznatků, byl navržen regulativní rámec pro velká data (kap. 5).

Dopad do praxe

Práce má vcelku přímočarý možný dopad do praxe v tom, že popisuje a vymezuje etické aspekty práce s velkými daty, což je v době masivního nárůstu množství dat generovaných chytrými zařízeními a infrastrukturami klíčový trend. Tento trend se dále posílí s rostoucím využitím metod umělé inteligence, kde bude zcela nezbytné mnoho eticky-relevantních rozhodnutí udělat již ve fáze koncipování obchodních příležitostí a záměrů generování a využití takových dat. Bez tohoto ohledu to nebude možné. Práce nabízí první ucelený rámec, jak etické aspekty na základě teoreticky fundovaných postupů popsat a uchopit.

Přínos doktoranda

Doktorand prokázal schopnosti kombinovat výzkumná paradigmata do výsledné výzkumné práce. V českém prostředí zpracoval původní průzkumy vstupující do regulatorního rámce pro velká data ve smyslu uživatelského vnímání závažnosti, přínosů a rizik masivního zpracovávání specifických velkých dat s vazbou na uživatele, zejména z telekomunikací. Pohled uživatelů jakožto primárních původců dat klíčovým způsobem vymezuje jejich další aplikovatelnost a rovněž úzce souvisí s „digital divide“ a informační asymetrií/nerovnováhou mezi možnostmi, vhledem a zájmy „velkého byznysu“, případně státních aktérů vs možnosti běžných koncových uživatelů. Zde vidím další přínos doktoranda.

Rozsah a kvalita rešerše

Doktorand jak v úvodních teoretických sekcích, tak v průběhu celého výzkumu tak, jak je v textu práce zachycen, do hloubky pracoval s dostupnou literaturou na vysoké úrovni. Zejména hodnotím, že byl schopen pracovat s literaturou v různorodých oblastech – filozoficko-etická, sociologická, právní, informatická, z oblasti řízení.

Faktické chyby

Spíše se jedná ne o chyby, ale o některé opomenuté zejména legislativní dokumenty a fakta související se zcela aktuálním tématem, a to je budoucí masivní využití umělé inteligence.

V práci není zmínka o ePrivacy směrnici/nařízení, což je lex specialis ke GDPR a věnuje se ochraně osobních údajů v elektronických komunikacích. Tato právní úprava se např. věnuje sběru údajů z datového provozu, cookies, apod., což je z pohledu big data důležité a je to přesně v předmětné oblasti zájmu – data z telekomunikačních systémů.

V práci se rovněž nediskutuje možnost “zosobnění” dat, tj. postup, kdy pro dataset, jenž neobsahuje v dnešním chápání osobní údaje, se může stát, že při nasazení ještě neexistující technologie (např. AI) nebo rozšíření o další (doposud neuvažované) datasety, se data mohou stát osobními, tj. bude možné identifikovat s údaji konkrétní osoby. To je teď v souvislosti s big data a GDPR široce diskutovaná otázka.

Úroveň zpracování textu

Text je zpracován v angličtině na solidní úrovni, stejně jako většina průběžných publikačních výstupů.

Celková vizuální a typografická stránka je na slušné úrovni, tabulky, seznamy a grafy čitelné, většina pochází od autora či spoluautorského týmu, původ je řádně uveden.

Publikace

Autor disertace má za sebou rozsáhlou publikační činnost rozprostřenou jak časově, tak tematicky do širšího záběru a podloženou dvacetiletou praxí na významných pozicích v sektoru telekomunikací. Výrazná většina prací je konferenčního charakteru na tuzemských a evropských fórech (indexované SCOPUS, CPCI). Jedna publikace je časopisecká (čas. Systémová integrace). Tematicky se jedná o široké pole témat mezi sociálními sítěmi a médii přes velká data, chování ve virtuálním prostředí a využití moderních médií v praxi velkých organizací a nově rovněž Industry 4.0. Celkově tvoří publikované práce sice široký rámec, ale s vcelku viditelným překryvem s tématem disertace a s přínosností pro disertační práci.

Ohlasy

Ve SCOPUS vykazuje t.č. 8 dokumentů a 8 citací, h-index 2 (2014-2018 se zřetelnou dynamikou růstu v posledních letech – 2018). Pro lepší zmapování ohlasů by bylo vhodné, kdyby autor měl registrovaný profil rovněž na Google Scholar. Považuji počty ohlasů na práce doktoranda za vyhovující.

Závěr

Práci považuji za kvalitní, tematicky pionýrskou, psanou se znatelnou zkušeností v oblasti velkých dat, jejich řízení, získávání znalostí a interpretace. S ohledem na výše uvedené skutečnosti práci doporučuji k obhajobě před příslušnou zkušební komisí pro obhajoby disertačních prací. Při obhajobě prosím autora o zodpovězení následujícího doplňujícího dotazu:

- *Kde vidíte hlavní etické dopady nového evropského nařízení o respektování soukromého života a ochraně osobních údajů v elektronických komunikacích a o zrušení směrnice 2002/58/ES (nařízení o soukromí a elektronických komunikacích), také známé jako nařízení ePrivacy?*

Brno, 9. ledna 2020

Doc. RNDr. Tomáš Pitner, Ph.D.

Vedoucí laboratoře Lasaris
Masarykova univerzita
Fakulta informatiky
Botanická 68a
602 00 Brno

E-mail: 94@muni.cz

Tel. 778884996