

UNIVERZITA KARLOVA

Filozofická fakulta

Katedra psychologie



BAKALÁŘSKÁ PRÁCE

Bc. Petra Fousková

**Nové psychometrické přístupy k měření osobnosti –
analýza digitálních stop a výsledků činnosti**

**New psychometric approaches to personality
measurement – digital footprints analysis**

Praha 2019

Vedoucí práce: PhDr. Eva Höschlová, Ph.D.

Ráda bych tímto poděkovala PhDr. Evě Höschlové, Ph.D., za cenné rady, které mi během psaní této práce poskytla. Mé poděkování také patří rodině a přátelům, kteří mi byli a jsou velkou oporou.

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 15. 5. 2019

.....

Bc. Petra Fousková

Abstrakt

Digitální technologie se stávají neodmyslitelnou součástí našich profesních i osobních životů. Posíláme e-maily, sdílíme příspěvky na sociálních sítích, pořizujeme selfie fotografie. To přináší nové příležitosti pro využití jiného typu dat, než bylo dříve pro měření osobnosti v psychometrii typické. Předkládaná práce je zaměřená na možnosti využití digitálních stop a výsledků činnosti pro odhad osobnostních vlastností a inteligence. Literárně přehledová část se zabývá možnostmi měření, resp. predikce osobnostních vlastností pomocí analýzy textu, fotografií a uživatelského chování na internetu. Návrh výzkumu je vystavěn nad otázkou, s jakou přesností je možné predikovat osobnostní rysy měřené pomocí dotazníku HPI na základě analýzy stop a výsledků činnosti na profesní sociální síti LinkedIn. Poznatky z navrhovaného výzkumu mohou být potenciálně přínosné jako jeden zdroj informací při náboru či výběru zaměstnanců.

Klíčová slova

psychometrie, digitální stopa, odhad osobnosti, sociální sítě, LinkedIn

Abstract

Digital technologies have become an inherent part of both our personal and professional lives. We send e-mails, share posts on social networks, take selfies. This has been bringing new opportunities for using different type of data than it was typical for psychometry in the past. This thesis deals with the possibilities of using digital footprints for personality traits and intelligence measurement. The first part of the thesis focuses on possibilities of personality traits estimation, or rather prediction, based on an analysis of text, photography and online user behaviour. The second part examines the extent to which is possible to estimate personality traits measured by Hogan Personal Inventory from digital footprints on the professional social network LinkedIn. The outcome from the proposed research could be potentially used as a source of information in the employee selection process.

Keywords

Psychometry, Digital Footprint, Personality Computing, Social Networks, LinkedIn

Obsah

Obsah	6
Úvod.....	7
Literárně přehledová část.....	9
1 Analýza textu	10
1.1 Pennebakerovy výzkumy	11
1.2 Další studie.....	14
1.3 Výhody a limity analýzy textu	16
2 Analýza fotografií	17
2.1 Odhad osobnostních rysů	18
2.2 Odhad inteligence.....	21
3 Analýza aktivity na sociálních sítích	23
3.1 Osobní sociální sítě	24
3.2 Profesní sociální sítě	26
Návrh výzkumného projektu	29
4 Výzkumný problém	29
5 Design výzkumného projektu	29
5.1 Metody získávání dat	32
5.2 Hypotézy	34
5.3 Výzkumný soubor	36
5.4 Metody zpracování a analýzy dat.....	38
6 Etika výzkumu	39
7 Diskuse.....	39
Závěr	44
Seznam použité literatury	45
Seznam tabulek	53
Seznam použitých zkratk	54
Příloha 1	I
Příloha 2	II
Příloha 3	III

Úvod

Tématem předkládané bakalářské práce jsou nové psychometrické přístupy k měření osobnosti, konkrétně se zaměřením na nové zdroje informací o člověku, kterými jsou digitální stopy a výsledky činnosti. Práce si klade za cíl na základě odborné literatury, článků a studií předložit ucelený pohled na možnosti využití digitálních stop k odhadu osobnostních vlastností a inteligence. Empirická část následně představuje návrh výzkumu, který by dosavadní studie daného tématu obohatil o poznatky zcela nové. Výběr tématu byl podpořen mojí osobní zkušeností z práce v personální agentuře, která se specializuje na nové přístupy k náboru zaměstnanců. Mezi ty patří například oslovování kandidátů prostřednictvím sociálních sítí. Začala jsem si pokládat otázku, zda je možné z profilu kandidátů na LinkedInu a dalších sociálních sítích vyvozovat validní závěry o jejich osobnosti. Hlubší pochopení této problematiky tak může přispět k zefektivnění náborových a výběrových procesů, stejně tak jako k obohacení tradičních psychometrických postupů.

První kapitola této práce se zabývá analýzou textu a je rozdělená do tří částí. První část je věnována výzkumům a projektům realizovaným Pennebakerem. Tyto projekty (např. LIWC, AnalyzeWords a další) jsou zástupci tzv. uzavřených přístupů k analýze textu a vycházejí zejména z analýzy funkčních slov. V druhé části první kapitoly je popsán výzkum založený na sběru dat prostřednictvím aplikace MoodPrism. Dále je zde představen český projekt CPACT, jehož výstupy byly zveřejněny teprve před několika měsíci – v listopadu roku 2018. V poslední části první kapitoly jsou shrnuty výhody a limity analýzy textu.

Druhá kapitola je zaměřená na analýzu fotografií. Profilové fotografie jsou dobře dostupným zdrojem informací, neboť jsou na většině sociálních sítích (např. Facebook, Twitter a Instagram) veřejně dostupné. Řada výzkumů se zabývá možnostmi odhadu osobnostních rysů z fotografií dle pojetí Big Five (např. Liu, Preoțiu-Pietro, Samani, Moghaddam, & Ungar, 2016; Kaurin, Heil, Wessa, Egloff, & Hirschmueller, 2018; Segalin et al., 2017). Jiné výzkumy jsou zaměřené na odhad inteligence na základě fotografií (např. Wei & Stillwell, 2016; Kleisner, Chvátalová, & Flegr, 2014).

Třetí kapitola vychází z výzkumů zabývajících se odhadem osobnosti uživatelů sociálních sítí na základě jejich aktivity na těchto sítích. Tematicky je třetí kapitola rozdělená na dvě části. První část popisuje výzkumy o analýze aktivity na osobních sociálních sítích, zejména na Facebooku (Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011; Youyou, Kosinski, & Stillwell, 2015; Yarkoni, 2010). Druhá část je věnována studiím zaměřeným na profesní sociální sítě, například LinkedIn (Ma & Leung, 2018; van de Ven, Bogaert, Serlie, Brandt, & Denissen, 2017; Rastogi & Nagappan, 2016).

Návrh vlastního výzkumu zkoumá možnosti odhadu osobnostních rysů uživatelů LinkedInu v sedmifaktorovém pojetí podle Hoganova osobnostního dotazníku, a to na základě jejich chování na LinkedInu a podoby jejich LinkedIn profilu. Tato část práce obsahuje také informaci o navrhovaném způsobu sběru a analýzy dat, diskusi etických aspektů, i diskusi možných výsledků a limitů zvoleného designu studie.

Seznam použitých zdrojů je ovlivněn tím, že se dané téma dostává do středu zájmu výzkumníků až od druhé dekády tohoto tisíciletí. Například v roce 2011 uvedli Gosling, Augustine, Vazire, Holtzman a Gaddis, že se jen málo psychologických výzkumů zaměřuje na projevy osobnosti na sociálních sítích. Dnes je situace opačná. Ke zpracování práce bylo tedy k dispozici dostatečné množství elektronických článků a studií, pocházejících zejména z Velké Británie a Spojených států amerických (např. Pennebaker, 2011b; Kosinski, Bachrach, Kohli, Stillwell, & Graepel, 2014; Yarkoni, 2010). Oproti tomu tištěných monografií je v této oblasti zatím k dispozici minimum. Výjimku tvoří publikace s názvem *The Secret Life of Pronouns: What Our Words Say About Us* od Pennebaker (2011a). Stejně tak i tuzemských zdrojů je poskrovnu, s výjimkou několika studií. V této oblasti stojí za zmínku projekt CPACT z Jihočeské univerzity (Kučera et al., 2018) a studie autorů Kleisnera, Chvátalové a Flegra (2014) z Karlovy Univerzity. V práci je citováno podle normy APA (2010).

Literárně přehledová část

Vývoj přístupů k měření osobnosti je ve 21. století výrazně ovlivněn rozmachem digitálních technologií, které poskytují v rámci psychometrie nové možnosti a zdroje dat. Na světě je více než 4 000 000 000 uživatelů internetu a jejich počet neustále roste. Každou sekundu je posláno kolem 2 800 000 e-mailů, sdíleno před 50 000 statusů na Facebooku a přidáno kolem 6 000 příspěvků na Twitteru (tzv. tweetů). Pro porovnání – v roce 2001 mělo připojení na internet jen kolem 14 000 000 lidí, což je zhruba 285krát méně, než jak je tomu dnes („Internet Live Stats”). Jedná se tedy o obrovské množství dat, které může být v rámci moderních přístupů k psychometrii využito k měření osobnosti.

Výzkumy zaměřené na odhad osobnostních vlastností a inteligence na základě digitální stopy uživatelů sociálních sítí lze na obecné úrovni rozdělit podle předmětu zájmu do dvou kategorií. Do první kategorie spadají výzkumy zabývající se tím, jak lidé hodnotí uživatele sociálních sítí na základě jejich aktivity na těchto sítích. Tyto výzkumy zkoumají například to, jaký vliv má počet přátel či výběr tzv. selfie fotografie jako profilové fotografie na úsudek probandů o osobnosti uživatelů sociálních sítí. Těchto výzkumů bylo realizováno nespočet (např. Utz, 2010; Evans, Gosling, & Carroll, 2008; Pelled et al., 2017). Úsudek probandů však v tomto typu výzkumů není porovnán s vnějším kritériem ve smyslu již validizovaného osobnostního dotazníku nebo jiné psychodiagnostické metody. To, že probandi vnímají uživatele sociální sítě jako extravertního, neboť má hodně přátel, kteří na fotografiích působí také extravertně, ještě neznamená, že posuzovaný uživatel sociální sítě skutečně extravertní je. Jedná se o oblast laického usuzování, a nikoliv o psychodiagnostiku, proto se touto oblastí předkládaná práce vzhledem ke svému rozsahu nebude zabývat. Do druhé kategorie výzkumů pak spadají studie zaměřené na možnosti využití digitálních stop a výsledků činnosti jako vstupního zdroje dat pro psychometrické testování. A právě tímto typem studií se zabývá tato kapitola.

Mezi nejrobustnější výzkumy čerpající z dat ze sociálních sítí patří projekt MyPersonality Stillwella a Kosinskeho, jehož počátky sahají do roku 2007. Za účelem výzkumu vytvořili autoři tohoto projektu facebookovou aplikaci, která

nabízela uživatelům Facebooku možnost vyplnění osobnostního dotazníku, který uživatelům následně poskytl zpětnou vazbu k dosaženému výsledku. Nad rámec toho mohli uživatelé aplikaci MyPersonality poskytnout informace ze svého facebookového profilu, k čemuž dalo svolení zhruba 40 % z nich. Tímto způsobem byly získány informace od 6 000 000 lidí, což představuje unikátní soubor dat, na jehož základě vznikla řada vědeckých studií (např. Kosinski, Wang, Lakkaraju, & Leskovec, 2016; Lambiotte & Kosinski, 2014; Cantador, Fernández-Tobías, & Bellogín, 2013; a mnoho dalších). Kromě vědeckých účelů byly výstupy projektu MyPersonality využity i ke zkvalitnění legislativy v oblasti ochrany soukromí. Udržování aplikace v provozu bylo ale vzhledem k časové náročnosti v roce 2012 ukončeno a od roku 2018 přestali tvůrci poskytovat získaná data jiným výzkumníkům, opět kvůli časové náročnosti („MyPersonality.org“).

1 Analýza textu

Po několik desetiletí byly za zlatý standard v přístupech k měření osobnosti považovány osobnostní dotazníky, neboť jsou rychlé, finančně dostupné a dosahují dobrých testových vlastností – tedy objektivity, standardizace, reliability a validity (Boyd & Pennebaker, 2017). Jak ovšem upozorňují Boyd a Pennebaker (2017), osobnostní dotazníky mají i svou nevýhodu. Ta spočívá zejména v tom, že se jejich výstupy zpravidla odvíjí od toho, co o sobě testované osoby vypoví, tedy na vnímané identitě probandů. Tento nedostatek překonává počítačová textová analýza. Mezi její výhody patří možnost zpracovat v relativně krátkém čase velké množství dat, získaných například se souhlasem majitelů z mobilních telefonů a sociálních sítí. Tato data mají navíc vysokou ekologickou validitu.

Výzkumy založené na počítačové analýze textu lze rozdělit do dvou skupin, a sice na přístupy uzavřené (closed approaches) a přístupy otevřené (open approaches). Oba typy přístupů mají své výhody i limity. Častější jsou výzkumy spadající do kategorie uzavřených přístupů (Kučera et al., 2018), které vychází z frekvence zastoupení jednotlivých lingvistických kategorií. Příkladem je dílo Pennebaker (viz podkapitola 1.1), či tuzemský projekt CPACT (popis projektu je v podkapitole 1.2). Proto je tomuto typu výzkumů v rámci této práce věnován větší

prostor. Oproti tomu otevřené přístupy nemají a priori definované kategorie, ale vychází pouze z analyzovaného textu. Příkladem otevřeného přístupu ke komputační analýze textu je latentní sémantická analýza, popsaná ve Wolfe et al. (1998).

1.1 Pennebakerovy výzkumy

Snad nejznámějším současným světovým odborníkem na analýzu jazyka z psychologického hlediska je sociální psycholog a profesor z Texaské univerzity v Austinu, James W. Pennebaker. Pennebaker se ve své práci zaměřuje především na osobnostní determinanty užití jazyka a dále na vztah jazyka a psychického a fyzického zdraví („James W. Pennebaker“, c1996–2018). Ve své práci rozlišuje Pennebaker slova do dvou kategorií. Jedná se o:

- Slova obsahová, která jsou nositeli významu. Do této skupiny patří například podstatná jména, slovesa, přídavná jména a příslovce.
- Slova funkční, která sama o sobě význam nenesou. Obvykle jsou krátká a běžnému posluchači/čtenáři těžko postřehnutelná, neboť jim lidé přirozeně nevěnují mnoho pozornosti. Konkrétně se jedná o zájmena, členy, předložky, pomocná slovesa, záporny, spojky atd.

Svou pozornost Pannebaker zaměřuje především na druhou jmenovanou kategorii, kterou označuje s trochou nadsázky jako klíč do duše. Na základě analýzy používání funkčních slov je možné s vysokou mírou spolehlivosti usuzovat na psychologické a demografické charakteristiky, jako je např. gender, věk, sociální zařazení, emocionální stav, úroveň upřímnosti, schopnost vést lidi, arogance, sociální uzavřenost a kvalita interpersonálních vztahů (Pennebaker, 2011, „LIWC“). Každý člověk totiž při přirozené produkci jazyka volí z mnoha možností, nehledě na to, zda se jedná o projev psaný nebo mluvený. Mluvčí/pisatel se rozhoduje, co bude obsahem jeho sdělení a jakými slovy ho vyjádří, jakou použije větnou strukturu atd. Groom a Pennebaker (2005) to ilustrují na příkladu, kdy člověk při vyprávění o své rodině může na jednotlivé členy odkazovat prostřednictvím jejich sociálních rolí (otec, matka), vlastními jmény (Tomáš, Zuzana), zájmeny v první osobě (já, my), či zájmeny ve třetí osobě (on, ona).

Všechny tyto možnosti odkazují na rodinné příslušníky, zároveň se ale podle nich dá usuzovat na rozdílné kvality daného interpersonálního vztahu. Na základě analýzy užití funkčních slov lze usuzovat i na aktuální psychický stav, což ilustruje následující příklad. Při použití zájmena „já“ se lidé krátce soustředí na sebe. Ve vyšší míře používají zájmeno já lidé, kteří čelí zvýšenému psychickému či fyzickému tlaku, neboť je tento tlak vede k zaměření se na sebe sama. Zvýšený výskyt zájmena „já“ tak může být projevem deprese, či stresu („LIWC“).

Odhady osobnosti na základě textu mohou být uskutečněny automatizovaně, například prostřednictvím programu AnalyzeWords. Tento program odhaduje osobnost člověka na základě jeho příspěvků na Twitteru. Stačí zadat uživatelské jméno konkrétní osoby a AnalyzeWords během vteřiny vyhodnotí emocionální a sociální styl dané osoby a také její styl uvažování, a to včetně vysvětlení, jak program k danému výsledku dospěl. Ukázka výstupu AnalyzeWords je uvedena v Příloze 1.

AnalyzeWords je založený na počítačovém programu LIWC (Linguistic Inquiry and Word Count), jehož aktuální verze nese označení LIWC2015 („LIWC“). Za nápadem vytvořit program LIWC stojí Pennebaker, programátorsky je pod ním podepsaná Pennebakerova studentka Francis. Při analýze textu postupuje LIWC tím způsobem, že v první fázi spočítá všechna slova a následně je rozřadí do více než 70 lingvistických kategorií podle toho, zda se jedná o podstatná jména, slovesa, slovesa v minulém čase, pomocná slovesa, slova popisující emoce atd. V dalším kroku LIWC vypočte procentuální obsazení jednotlivých mluvnických kategorií a na základě toho vyvodí úsudek o analyzované osobě (Pennebaker, 2011a).

LIWC dosahuje dobrých psychometrických vlastností, jako jsou test-retestová validita, externí validita a vnitřní konzistence (Pennebaker & King, 1999). Mezi jeho hlavní výhody patří také rychlost, s jakou dokáže zpracovat i velké množství textu. Například analýza kompletního díla od Williama Shakespeara by zabrala méně než 20 sekund. Mezi nevýhody oproti lidskému hodnotiteli pak patří zejména to, že program nedokáže detekovat sarkasmus a humor. Krom toho od sebe nedokáže rozeznat homonyma, tedy slova, která se stejně píšou/vyslovují, ale mají

odlišný význam (Pennebaker, 2011a). Za zmínku stojí fakt, že LIWC je na rozdíl od projektu MyPersonality využíván nejenom k vědeckým, ale i komerčním účelům.

Analýzu textu prostřednictvím metody LIWC realizovali Pennebaker a King (1999). Tito autoři analyzovali texty 842 probandů a dospěli k závěru, že neuroticismus a extraverte pozitivně korelují s použitím negativně zabarvených emočních slov. Neuroticismus navíc negativně koreluje s výskytem pozitivně zabarvených emočních slov. Dále byla nalezena pozitivní korelace mezi extravertí a odkazováním se v textu na ostatní lidi. Přívětivost dle výsledků této studie pozitivně koreluje s pozitivními emocemi a negativně koreluje s negativními emocemi.

Metodu LIWC lze kromě odhadu osobnostních vlastností a inteligence použít i k predikci dalších jevů, jako je například úspěšnost vysokoškolského studia (Pennebaker, Chung, Frazee, Lavergne, & Beaver, 2014), či proměny psychického stavu v závislosti na prožívané tragické události (Cohn, Mehl, & Pennebaker, 2004). Studie zaměřující se na tato témata však souvisí s tématem předkládané práce spíše okrajově, proto budou pro zajímavost uvedeny jen dvě.

Baddeley, Danielová a Pennebaker (2011) analyzovali texty anglického cestovatele a průzkumníka Henryho Hellyera, který ve věku 42 let spáchal sebevraždu. Výzkumný tým analyzoval knihy, dopisy a deníky, které Hellyer napsal v posledních 7 letech svého života. Analýza odhalila, že se v textech zvyšovala četnost výskytu zájmena v první osobě jednotného čísla („já“), naopak výskyt zájmen v první osobě množného čísla („my“) se snižoval. Také se zvyšoval výskyt slov s negativním emočním zabarvením. Výzkumný tým na základě toho dospěl k závěru, že Hellyer v době před spácháním sebevraždy více zaměřoval pozornost sám na sebe, byl více sociálně izolovaný a převažovaly u něj negativní emoce. Autoři této studie však upozorňují, že se jedná o ojedinělou kazuistiku, proto je třeba opatrnosti při generalizaci výsledků. Tento limit však překonává výzkum Stirmana a Pennebaker (2001), v jehož rámci bylo analyzováno 300 básní od 18 básníků, z čehož 9 básníků spáchalo sebevraždu. Stirman a Pennebaker objevili rozdíly v jazykovém projevu obou skupin. Projev básníků, kteří sebevraždu

spáchali, obsahoval více slov týkajících se jich samotných, a naopak méně slov souvisejících s ostatními lidmi.

Další studie vedla Pennebaker k závěru, že lidé, kteří prožili traumatickou událost a tento prožitek si nechali pro sebe, mají více zdravotních obtíží než lidé, kteří také prožili traumatickou událost, ale otevřeně o ní mluví. I zde využil Pennebaker počítačnické jazykové analýzy, konkrétně analyzoval psaný projev traumatizovaných lidí. Všiml si, že čím byl jejich zdravotní stav lepší, tím více přecházeli od používání zájmen v první osobě jednotného čísla (já, mě, moje) k ostatním zájmenům (ona, my, ty) (Pennebaker, 2011a).

Na co se naopak na základě analýzy jazyka zřejmě usuzovat nedá, je sexuální orientace. Domněnku, že se homosexuální muži lingvisticky odlišují od mužů heterosexuálních, a homosexuální ženy od žen heterosexuálních, nepodpořil výzkum Grooma a Pennebaker (2005). V rámci tohoto výzkumu byla provedena analýza 1 500 příspěvků z internetových seznamovacích portálů od heterosexuálních i homosexuálních mužů a žen, která ale neobjevila statisticky významné rozdíly v jazykovém projevu těchto skupin.

1.2 Další studie

Kromě Pennebaker se i řada dalších autorů zabývala otázkou, zda je možné na základě četnosti výskytu slov vyjadřujících pozitivní či negativní emoce usuzovat na osobnostní vlastnosti, či zda je za pomoci výskytu určitých slov možné predikovat konkrétní psychické jevy. Kern, Rickard, Seabrook a Fulcher (2018) uvádí, že takové studie se zpravidla omezují na analýzu frekvence výskytu klíčových slov, což může být na škodu, neboť psychický stav je reflektován ve více aspektech než jen v poměru zastoupení jednotlivých lingvistických kategorií. Tito autoři se zaměřují na vývoj variability užití emočně zabarvených slov v čase. Konkrétně zkoumali souvislost mezi vývojem užití pozitivně a negativně zabarvených emočních slov ve statusech na Facebooku a Twitteru a pocíťovanými depresivními symptomy. Depresivní symptomy měřili pomocí mobilní aplikace MoodPrism, která vznikla za podpory Univerzity v Melbourne a Monashské Univerzity. MoodPrism umožňuje uživatelům pravidelně zaznamenávat odpověď

na krátké výroky (např. „dnes se cítím aktivně“ či „mám ze sebe dobrý pocit“) a v závislosti na tom poskytuje informaci o momentálním psychickém stavu uživatelů. Uživatelé mohou také zaznamenávat, s kým se ve chvíli vyplňování hodnocení nacházeli a jakou činnost vykonávali. Na základě tohoto hodnocení jim pomáhá MoodPrism odhalit vzorce chování, které vedou k pocíťování pohody, nebo naopak úzkosti a napětí. Výstupy této studie naznačují, že nestálost užívání negativně zaměřených emočních slov ve statusech na Facebooku může sloužit jako prediktor prožívané deprese. Zajímavé ale je, že v případě Twitteru došel výzkum k opačnému výsledku – tedy že větší variabilita v užívání slov popisujících negativní emoce v tweetech predikuje nižší míru depresivních symptomů. Výstup z této studie tedy tvrdí, že analýza užití slov označujících negativní emoce může být smysluplně využita jako prediktor deprese, ovšem musí být brán v potaz typ sociální sítě, která slouží jako zdroj dat.

Co se analýzy českého jazyka týče, v nedávné době byl realizován na Pedagogické fakultě Jihočeské univerzity v Českých Budějovicích unikátní projekt s názvem CPACT (z anglického Computational Psycholinguistic Analysis of Czech Text nebo česky komputační psycholinguvistická analýza českého textu). Jedná se o první výzkum tohoto typu, který vychází z českého jazyka. Zároveň jde o novinku na poli české psycholinguistiky, neboť sběr dat probíhal v letech 2016 a 2017 a výsledky byly s veřejností sdíleny teprve v listopadu roku 2018. CPACT se zaměřil: „*na studium souvislosti mezi osobností člověka a slovy, která užívá; zejména na úrovni vztahů lingvistických charakteristik psaného a mluveného textu s výstupy psychologických testů zaměřených na sebeposouzení jedince a jeho posouzení druhou osobou*“ (Kučera et al., 2018, s. 4). Kromě pilotáže probíhal výzkum ve třech modulech. Prvního modulu se zúčastnilo 200 participantů z běžné populace, kteří byli vybráni kvótním výběrem. Druhého modulu se zúčastnilo 72 osob z klinické populace, které byly diagnostikovány s úzkostnou a/nebo depresivní symptomatikou. V rámci třetího modulu pak zaškolení hodnotitelé laikové (ne-psychologové) posuzovali všechny získané texty na základě předem stanovených kritérií. Krom toho byly získané texty podrobeny lingvistické analýze, která sledovala u každého slova, věty i textu samotného určité proměnné (např. emoční náboj, slovní druh, výskyt vulgárních slov atd.). Výsledky studie

mimo jiné podpořily myšlenku, že odhadovat osobnost člověka na základě jeho jazykového projevu v českém jazyce je smysluplné. Konkrétně v kontextu modelu Big Five dospěla studie k následujícím závěrům:

- **Otevřenost vůči zkušenosti:** čím více byli lidé otevření vůči zkušenosti, tím rozsáhlejší byla jejich slovní zásoba a také častěji používali neutrální emoční slova (slova, která nelze označit za pozitivní, ani za negativní). Zároveň tito lidé používali méně interpunkce.
- **Svědomitost:** signifikantní vztahy u rysu svědomitosti byly zjištěny jen u textů neformálních (probandi v rámci výzkumu tvořili různé typy textů). Rys svědomitosti negativně koreloval s výskytem výrazů v negativním tvaru a pozitivně koreloval s výskytem neurčitých číslovek, podstatných jmen, přídavných jmen a přivlastňovacích zájmen.
- **Extraverze:** co se rysu extraverze týče, zde byly signifikantní vztahy nalezeny pouze u mluvených textů. Konkrétně extraverze pozitivně korelovala s počtem zájmen a výskytem přičestí minulého.
- **Přívětivost:** pro rys přívětivosti bylo sice nalezeno několik vztahů, ovšem po korekci žádný z nich nedosáhl požadované hodnoty $p < 0,05$.
- **Neuroticismus:** tento rys koreloval pozitivně s výskytem zájmen v první osobě a sloves a negativně s výskytem neurčitých číslovek a bohatostí slovní zásoby.

1.3 Výhody a limity analýzy textu

Kromě výše popsaných zdrojů mohou jako vstupní data pro analýzu textu sloužit i facebookové statusy (Schwartz et al., 2013), e-mailová komunikace (Gill & Oberlander, 2003), či dokonce e-mailové adresy (Back, Schmukle, & Egloff, 2008). Yarkoni (2010) však upozorňuje na tři omezení, se kterými se výzkumy zaměřené na odhad osobnostních charakteristik na základě lingvistické analýzy

řeči/textu potýkají. Za prvé, tyto studie často analyzují ústní či písemné projevy, které vznikly v laboratořích. V rámci výzkumů jsou participanti obvykle vyzváni, aby hovořili nebo psali na zadané téma a nejedná se tedy o spontánní a přirozený projev. Zůstává tedy otázkou, na kolik lze výsledky těchto studií generalizovat i na běžný život, v jehož rámci si lidé nevybírají, pouze jak budou hovořit/psát, ale zejména o čem. Za druhé, záznamy řeči / texty jednotlivých probandů jsou obvykle kvůli organizačním důvodům rozsahově limitovány na pár tisíců slov. Navíc jsou data často sbírána jednorázově, což neumožňuje vyvozovat závěry ohledně stability a vývoje užití jazyka v čase. A za třetí, většina studií využívá k popisu osobnosti poměrně široké modely, jako například Big Five, a nezaměřují se na specifitější a užší osobnostní dimenze.

Na seznam nevýhod a omezení v generalizaci výsledků studií zaměřených na analýzu jazyka lze přidat také fakt, že většina studií vychází z analýzy anglického jazyka. Zajisté je tedy na místě ptát se, zda jsou výsledky těchto studií zobecnitelné i na projev v jiných jazycích. Odpověď na tuto otázku hledali Qiu et al. (2017) v rámci výzkumu zkoumajícího vztah osobnosti a užití čínského jazyka. Dospěli k závěru, že některé aspekty odhadu osobnosti na základě jazyka jsou univerzální a platné pro mluvčí v anglickém i čínském jazyce, jiné aspekty jsou naopak kulturně specifické. Na základě tohoto zjištění doporučují přístup k analýze jazyka založený na rozboru slov nesoucích obsah, spíše než na slovech funkčních, neboť funkční slova jsou dle těchto autorů citlivější na kulturní specifitnost.

Výstupy analýzy textu mají v praxi širokou škálu uplatnění. Kromě výše zmíněných odhadů osobnostních rysů, školní úspěšnosti, detekce deprese a sebevražedného chování je možné odhadovat i demografické charakteristiky (věk, pohlaví a bydliště) (Ikeda, Hattori, Ono, Asoh, & Higashino, 2013), výsledky analýz tak mohou být využity například pro marketingové účely v rámci přesnějšího cílení obsahu na konkrétní skupiny uživatelů.

2 Analýza fotografií

Lidé žijící v moderní společnosti si často pořizují fotografie a tzv. selfies a následně je sdílejí prostřednictvím sociálních sítí. Některé fotografie umístěné na

osobních profilech uživatelů sociálních sítí jsou pro výzkumníky vzhledem k nastavení soukromí nepřístupné, ale například profilové fotografie jsou ve většině sociálních sítí veřejné (např. na Facebooku, Twitteru a Instagramu). Jak uvádí Wei a Stillwell (2016), profilové fotografie poskytují informaci ohledně chování (postoj, oblečení, lokalita, přítomnost dalších lidí na fotografii), i preferencí (přítomnost zvířat, sportovního vybavení, atributy společenského postavení) uživatelů sociálních sítí. To, jakou fotografii si lidé zvolí jako profilovou, je determinováno jejich osobností. Opět se tak jedná o nový typ zdroje dat, který může být využit k psychometrickým účelům.

2.1 Odhad osobnostních rysů

Segalin et al. (2017) zkoumali možnosti odhadu osobnostních rysů z facebookových profilových fotografií. V rámci výzkumu porovnávali automatické usuzování provedené počítačem na vzorku 11 736 profilových fotografií různých uživatelů s usuzováním probandů a došli k závěru, že počítač předčil lidský úsudek, co se odhadu extravertze a neuroticismu týče. Dále zjistili, že lidé skórující vysoko na škále extravertze a/nebo přívětivost často používají fotografie v teplých barvách. Tito lidé také měli tendenci vybírat profilové fotografie, které zachycují expresivní výrazy tváře, což patrně koresponduje s jejich tendencí se socializovat. Lidé skórující vysoko na škále neuroticismu mají dle tohoto výzkumu tendenci vybírat profilové fotografie, které jsou pořízené ve vnitřním prostředí.

Liu, Preoțiu-Pietro, Samani, Moghaddam a Ungar (2016) se v rámci své studie zaměřili na možnosti odhadu osobnostních rysů podle modelu Big Five na základě analýzy tweetů a profilových fotografií 66 000 uživatelů Twitteru. Co se odhadu osobnostních rysů na základě profilových fotografií týče, došli k následujícímu zjištění:

- **Otevřenost vůči zkušenosti:** lidé skórující vysoko na škále otevřenosti vůči zkušenosti často volili nekonformní fotografie, na kterých nebyly zachyceny žádné tváře. Zdá se, že kladli důraz na estetičnost fotografií – jejich fotografie byly kontrastní, ostré a nerozmazané. Na druhou

stranu byly tyto fotografie také méně barevné a nepůsobily autentickým dojmem. Co se doplňků týče, lidé otevření vůči zkušenosti byli často zachyceni s dioptrickými brýlemi. Na fotografiích projevovali oproti ostatním rysům častěji negativní emoce, zejména hněv. Na druhou stranu se méně usmívali a také méně projevovali pozitivní emoce, například radost.

- **Svědomitost:** pro lidi s vysokým skórem na škále svědomitosti patrně estetická stránka fotografie není tak důležitá, jako pro první zmiňovanou skupinu. Svědomití lidé často volili fotografie obsahující alespoň jednu tvář, méně často pak fotografie s více tvářemi či dokonce úplně bez tváří. V porovnání s lidmi otevřenými vůči zkušenosti v tomto případě tváře na fotografiích zabíraly menší část snímku a neměly žádné brýle (dioptrické ani sluneční). Barevné ladění bylo spíše světlé, přirozené a barevné. Svědomití lidé vyjadřovali v porovnání s ostatními čtyřmi skupinami nejvíce emocí. Převažovala však expresivnost emocí nad přirozeností. Nejčastěji byly zastoupeny pozitivní emoce.
- **Extraverze:** lidé skórující vysoko na škále extraverze měli v porovnání s lidmi skórujícími nízko na této škále nejbarevnější fotografie. Dále preferovali snímky, na kterých bylo zobrazeno více lidí. Na rozdíl od všech ostatních kategorií, extraverti na fotografiích vypadali jako mladší, než ve skutečnosti byli, případně se s mladšími lidmi fotili. Co se emocí týče, extravertní lidé ukazovali stejně pozitivní emoce, jako lidé svědomití, jen s menší intenzitou. Typicky neměli dioptrické brýle, jejichž nošení je asociováno spíše s introvertními jedinci.
- **Přívětivost:** lidé s vysokým skórem na škále přívětivosti volili profilové fotografie obsahující tváře. Obecně byly jejich snímky spíše plné různých prvků než jednoduché. Také byly barevné, ačkoliv méně než fotografie extravertů. Dále byly méně jasné a méně ostré a nepůsobily příliš esteticky. Co se výrazů tváře týče, i zde převažovaly pozitivní emoce.

- **Neuroticismus:** u fotografií lidí s vysokým skórem na škále neuroticismu byla nalezena negativní korelace s barevností. Obecně fotografie působily jednoduchým dojmem, což je typické i pro otevřenost vůči zkušenosti. Na rozdíl od tohoto rysu byly ale snímky neurotických lidí méně estetické. Dále se o fotografiích neurotických lidí dá říct, že na nich často chyběly tváře a pokud byly přítomny, byly poměrně velké a měly nasazené dioptrické brýle. Co se míry projevení pozitivních emocí týče, zde byla nalezena negativní korelace.

Kaurin, Heil, Wessa, Egloff, a Hirschmueller (2018) realizovali výzkum zaměřený na možnosti usuzování na narcismus, sebevědomí a pět rysů dle modelu Big Five na základě selfie fotografií. Narcismus a sebevědomí zvolili vzhledem k předpokladu, že tyto dva rysy determinují do značné míry způsob, jakým se lidé na sociálních sítích prezentují. Jako výhodu navrhovaného zdroje informací uvádí fakt, že selfie fotografie jsou snadno dostupné. V rámci tohoto výzkumu také porovnávali přesnost úsudků vyvozených na základě selfie fotografií ze sociálních sítí a úsudků vyvozených na základě fotografií vytvořených v laboratorních podmínkách. Dospěli k závěru, že oba typy úsudků se do značné míry shodují. Shoda mezi posuzovateli byla vyšší, pokud tito hodnotili pouze na základě selfie. Ovšem s výjimkou hodnocení dvou rysů – přívětivosti a sebevědomí. Jako nejjednodušší rys k odhadu na základě selfie fotografií se ukázala otevřenost vůči zkušenosti. Narcismus a extravertze byly na základě selfie vnímány jako vyšší a svědomitost byla na základě selfie vnímána naopak jako nižší.

Nabízí se samozřejmě otázka, zda lidé na sociálních sítích prezentují své reálné, nebo ideální já. Segalin et al. (2017) zastávají názor, že uživatelé se na sociálních sítích nemohou prezentovat odlišným způsobem než mimo ně, neboť jsou online propojeni s lidmi, kteří je znají i v běžném životě. Tímto tématem se dále zabývali Vazire a Gosling (2004). Realizovali výzkum, v jehož rámci 11 probandů zhlédlo 89 osobních stránek různých uživatelů a na základě toho se pokusilo odhadnout osobnostní charakteristiky majitelů těchto stránek. Úsudek probandů byl následně porovnán se třemi kritérii: se sebehodnocením vlastníků internetových stránek, dále s jejich ideálním já a do třetice s hodnocením osoby

majiteli blízké. Výsledky výzkumu ukázaly na vysokou shodu mezi úsudky posuzovatelů. Také byla nalezena korelace mezi hodnocením posuzovatelů s vnějším kritériem (sebehodnocení, hodnocení známým člověkem). Na základě tohoto výzkumu tedy lze usuzovat, že informace získané z osobních stránek lze považovat za poměrně validní zdroj informací o osobnosti.

2.2 Odhad inteligence

Osobností rysy nejsou jedinou charakteristikou, na kterou se dá na základě fotografií usuzovat. Inteligence se však liší od osobnostních rysů v tom, že je všeobecně uznáváno, že je lepší být inteligentní více nežli méně. To v případě psychických rysů neplatí. Nelze bez kontextu tvrdit, zda je lepší být extravertní, či introvertní (Wei, & Stillwell, 2016). V kontextu výzkumů zaměřených na odhad inteligence je třeba rozlišit inteligenci vnímanou, danou úsudkem probandů, a inteligenci měřenou, danou výsledkem inteligenčních testů. Většina výzkumů na poli odhadu inteligence z fotografií se zaměřuje spíše na odhad vnímaní inteligence (např. Talamas, Mavor, & Perrett, 2016; Talamas, Mavor, Perrett, Sundelin, & Axelsson, 2016).

Wei a Stillwell (2016) zkoumali možnosti odhadu inteligence na základě profilových fotografií na Facebooku. Konkrétně pracovali se vzorkem 1 122 facebookových uživatelů, kteří jim poskytli své profilové fotografie a zároveň vyplnili 20položkový matricový inteligenční test, který byl validizovaný s Ravenovými progresivními matricemi. Výzkumníci z fotografií extrahovali jednotlivé charakteristiky pomocí metod automatické extrakce (například programem Face++ pro extrakci obličejových charakteristik). Tyto charakteristiky rozdělili do 4 oblastí – barva, kompozice, struktura a tělo a obličej. Odhad inteligence na základě fotografií byl taky učiněn 739 probandy – laiky. Tito odhadovali výši inteligence lidí na fotografiích a svůj odhad označovali na 7stupňové škále. Výsledky této studie naznačují následující:

- **Barva:** Byla nalezena pozitivní korelace mezi inteligencí měřenou inteligenčním testem a výskytem zelených pixelů. Oproti tomu růžové, fialové a červené pixely korelovaly s měřenou inteligencí negativně. Co

se vnímané (nikoliv měřené) inteligence týče, byla zde nalezena pozitivní korelace s výskytem šedých a bílých pixelů. Naopak hnědé, zelené, růžové a fialové pixely korelovaly s vnímanou inteligencí negativně.

- **Obličej a tělo:** Rozměry zobrazené kůže a obličeje negativně korelovaly s měřenou i vnímanou inteligencí. S vnímanou inteligencí navíc negativně korelovaly rozměry těla. Výzkum dospěl v této oblasti navíc k zajímavému zjištění, že pokud se osoba na fotografii směla nebo měla brýle, byla vnímaná jako inteligentnější. Jednalo se ovšem o falešné vodítko, neboť v tomto případě nebyla nalezena pozitivní korelace s inteligencí měřenou intelligenčním testem.
- **Struktura:** Byla zjištěna pozitivní korelace mezi ostrotí snímku a měřenou i vnímanou inteligencí. Naopak zrnitost korelovala s měřenou inteligencí negativně.
- **Kompozice:** Prvky týkající se kompozice nekorelovaly významně s měřenou ani vnímanou inteligencí. S výjimkou procenta okrajových pixelů (pixelů, které oddělují objekty od okolí). Ty korelovaly negativně s vnímanou inteligencí.

Kleisner, Chvátalová a Flegr (2014) zkoumali vliv tvaru tváře na vnímanou a na měřenou inteligenci. V rámci první části jejich výzkumu probandi hodnotili 40 fotografií s tvářemi žen a 40 fotografií s tvářemi mužů. Výzkumníci měli k dispozici nejen fotografie, ale i výsledky Testu struktury inteligence (IST 2000 R) těchto mužů a žen. Probandi, nehledě na jejich pohlaví, dokázali odhadnout inteligenci mužů, avšak odhady inteligence žen přesné nebyly. Cílem druhé části výzkumu těchto autorů bylo za pomoci geometrické morfometrie určit ty rysy obličeje, které mají vliv na vnímanou a/nebo na měřenou inteligenci. Dle výsledků tohoto výzkumu jsou jako inteligentní vnímány tváře, které jsou delší, mají větší nos, mírně zvednuté koutky úst a ostře řezanou bradu. Nebyla však nalezena korelace mezi tvarem tváře a měřenou inteligencí.

3 Analýza aktivity na sociálních sítích

V posledních dvou letech byly realizované dvě meta-analýzy, resp. dvě série meta-analýz, zaměřující se na odhad osobnostních rysů na základě digitálních stop a výsledků činnosti na sociálních sítích. Autoři první z nich jsou Azucar, Marengo a Settanni (2018). Tento výzkumný tým uskutečnil separátní meta-analýzu pro každý rys z pětifaktorového osobnostního modelu. Zaměřil se při tom na studie, které vycházejí z různých typů dat, jakými jsou například demografické údaje, profilové fotografie, jazyková analýza a aktivita na sociálních sítích (počet přátel, stránky označené jako „líbí se mi“ atd). Autoři dospěli k závěru, že průměrná přesnost odhadovaných korelací pro otevřenost vůči zkušenosti, svědomitost, extraverci, přívětivost a neuroticismus (v tomto pořadí) je $r = 0,39$; $r = 0,35$; $r = 0,40$; $r = 0,29$ a $r = 0,33$. Dále výzkumníci zjišťovali, do jaké míry ovlivňuje typ zdroje digitální stopy a výsledků činnosti přesnost odhadu osobnostních rysů. Také si položili otázku, jak se budou lišit odhady učiněné na základě zdrojů z různých sociálních sítí. Dospěli k zjištění, že s výjimkou přívětivosti lze docílit přesnějšího odhadu užitím dat z různých zdrojů. Odhady otevřenosti vůči zkušenosti, přívětivosti a neuroticismu byly přesnější, pokud byly vytvořeny na základě analýzy, která vycházela i z demografických údajů.

Druhou meta-analýzu realizovali Liu a Campbell (2017). Porovnávali přitom výstupy studií, které vycházejí z následujících informací ze sociálních sítí: způsob užívání, hraní her, vyhledávání informací, interakce s ostatními uživateli, sdílení fotografií a příspěvků, počet přátel. Z pětifaktorového modelu našli nejvíce statisticky významných vztahů mezi výše zmíněnými proměnnými a extraverci. Nejsilnější vztah ($r = 0,34$) byl nalezen mezi extraverci a počtem přátel. Také v případě otevřenosti vůči zkušenosti bylo nalezeno několik statisticky významných vztahů, například s hraním her ($r = 0,28$). Přívětivost korelovala se sdílením fotografií ($r = 0,18$), svědomitost s hraním her ($r = -0,22$) a neuroticismus se sdílením příspěvků ($r = 0,13$).

3.1 Osobní sociální síť

Sociální síť se dají rozdělit do dvou skupin, a sice na sociální síť osobní, které slouží k volnočasovým aktivitám a nejsou primárně používány k pracovním účelům, a profesní. Mezi osobní sociální síť patří například Facebook, YouTube, Twitter či VKontakte. Od tohoto typu je možné rozlišit profesní sociální síť, které jsou užívány za účelem navazování a udržování profesních kontaktů a posilování značky člověka či společnosti. Mezi profesní sociální síť je možné zařadit LinkedIn, ResearchGate, GitHub, Xinx a další.

Gosling, Augustine, Vazire, Holtzman a Gaddi (2011) realizovali výzkum zaměřený na odhad osobnosti na základě aktivity na Facebooku. V rámci jejich studie 159 probandů – univerzitních studentů – vyplnilo 10položkový osobnostní dotazník TIPI (Ten Item Personality Inventory), založený na pětifaktorovém modelu osobnosti (dotazník obsahuje dvě položky pro každý rys). Probandi dále odpovídali na 11 otázek týkajících se jejich aktivity a profilu na Facebooku:

1. počet přátel;
2. počet hodin strávených na Facebooku za týden;
3. frekvence, s jakou se dívají na různé stránky;
4. frekvence, s jakou se dívají na profily jiných uživatelů;
5. frekvence, s jakou komentují příspěvky ostatních;
6. frekvence, s jakou se dívají na vlastní profil;
7. frekvence, s jakou upravují informace na vlastním profilu;
8. frekvence, s jakou přidávají fotografie zobrazující pouze je;
9. frekvence, s jakou přidávají fotografie zobrazující je a zároveň další lidi;
10. frekvence, s jakou přidávají fotografie ostatních lidí, na kterých sami nejsou zachyceni;
11. frekvence, s jakou si mění profilovou fotografii.

Dle výsledků této studie využívají Facebook častěji extraverti než introverti. Facebook tak patrně představuje další možnost, jak mohou extraverti naplnit svoji potřebu sociálního kontaktu. Extraverti měli dále větší počet přátel, častěji psali komentáře na zdi přátel a také častěji aktualizovali informace o sobě. Lidé skórující

vysoko na škále přívětivosti častěji navštěvovali vlastní profil i profily ostatních uživatelů. Dále byla nalezena negativní korelace mezi škálou svědomitosti a časem stráveným na různých facebookových stránkách, což může být způsobeno tím, že je pro lidi s nízkým skórem na škále svědomitosti obtížnější se od Facebooku odtrhnout. Co se otevřenosti vůči zkušenosti týče, zde byla nalezena pozitivní korelace s přidáváním a nahrazováním fotografií. Pouze škála neuroticismu nekorelovala významně s žádnou z výše uvedených proměnných.

Youyou, Kosinski a Stillwell (2015) se zaměřili na odhad osobnostních rysů dle modelu Big Five na základě obsahu Facebooku, který respondenti označili, jako že se jim líbí. Funkcí „líbí se mi“, či anglicky „like“, jsou obecně uživateli označovány stránky, ke kterým mají pozitivní vztah. Může se jednat např. o hudební skupiny, produkty běžné spotřeby, sporty, restaurace atd. To, jaké stránky uživatelé označují jako „líbí se mi“, je determinováno jejich osobností. Dle autorů této studie mají například lidé s vysokým skórem na škále otevřenosti vůči zkušenosti tendenci označit jako „to se mi líbí“ stránky Salvadora Dalího, TED talks nebo meditace. Extravertním lidem se pak líbí stránky o tancování, večírcích a hvězdách show businessu. Autoři si položili otázku, zda je na základě analýzy „lajků“ schopný odhadnout osobnost lépe počítač, nebo lidé uživateli Facebooku blízcí. Za tímto účelem vyplnilo 86 220 respondentů 100 položkový osobností dotazník IPIP (International Personality Item Pool) a poskytlo výzkumníkům informace ze svých facebookových profilů. Výstupy studie naznačují, že počítač je schopný lépe odhadnout osobnost uživatelů Facebooku než osoby blízké. V některých aspektech odhady počítače dokonce předčily sebehodnocení. Konkrétně stačilo počítači 10 „lajků“, aby odhadl osobnost člověka lépe než jeho průměrný kolega. 70 „lajků“, aby předčil odhad spolubydlícího nebo kamaráda. Při 150 „lajcích“ si počítač vedl lépe než rodinní členové. A pokud měl počítač k dispozici 300 „lajků“, nevyrovnali se jeho odhadu ani manžel/manželka posuzované osoby. Úsudky vytvořené počítačem navíc dosahovaly vyšší externí validity v porovnání s lidmi při odhadu politického názoru a fyzického zdraví. Na základě výzkumu Kosinskeho, Stillwella a Graepely (2013) lze soudit, že facebookové „lajky“ mohou sloužit kromě odhadu osobnostních rysů i k predikci dalších charakteristik, mezi které patří např. věk,

pohlaví, inteligence, sexuální orientace, etnická a náboženská příslušnost, politické názory a užívání návykových látek.

Podobný výzkum realizoval tým ve složení Kosinski, Bachrach, Kohli, Stillwell a Graepel (2014). Tito autoři se zaměřili na možnosti odhadu osobnosti na základě analýzy webových stránek, které člověk navštívuje. Uvádí, že se jedná o první studii, která využívá tento zdroj informací k odhadu osobnostních rysů. Výzkumný tým pracoval se souborem dat získaných v rámci projektu MyPersonality. Měl k dispozici data od 350 000 lidí, kteří vyplnili IPIP. Účastníci výzkumu také vyplnili dotazník, který se jich dotazoval na frekvenci, s jakou navštěvují různé webové stránky. Třetí typ dat, který měli výzkumníci k dispozici, se týkal aktivity probandů na Facebooku. Konkrétně se jednalo o informace o obsahu, který účastníci označili jako „líbí se mi“, počet přátel, počet skupin, jichž byli probandi členi, počet událostí, kterých se probandi účastnili, počet fotografií, které probandi sdíleli a/nebo na nich byli označení atd. Dle výstupů této studie extraverti navštěvují signifikantně častěji webové stránky o muzice a sociálních sítích, zatímco introverti preferují spíše webové stránky věnované komiksům, filmům a literatuře. Webové stránky z oblasti psychického zdraví jsou častěji navštěvovány lidmi skórujícími vysoko na škále přívětivosti a/nebo svědomitosti.

Další informace o tomto tématu obsahuje komparativní analýza od Farnadiho et al. (2016), která porovnává možnosti užití Twitteru, Facebooku a YouTube k odhadu osobnosti.

3.2 Profesionální sociální síť

Van de Ven, Bogaert, Serlie, Brandt a Denissen (2017) uvádí několik oblastí, ve kterých se liší profesionální sociální síť od těch osobních. Podle těchto autorů lidé přidávají obsah na profesionální sociální síť s větším rozmyslem než na osobní sociální síť, neboť si jsou vědomi toho, že obsah mohou vyhledat jejich současní či potenciální zaměstnavatelé. To činí profesionální sociální síť více statickými. Dále je na nich v porovnání s osobními sociálními sítěmi obvykle méně informací, neboť je více limitován rozsah profilů/příspěvků. Krom toho, jak ukazuje výzkum Garcii a Sikströma (2016), lidé mají tendenci používat jiná slova, pokud mají popsat svoji

osobnost přátelům, oproti situaci, ve které mají popsat svoji osobnost potenciálnímu zaměstnavateli. Pokud o sobě mluví mezi přáteli, používají vřele působící slova, jako např. laskavý či milující. Na druhou stranu, pokud se mají popsat v rámci výběrového řízení, použijí spíše slova označující kompetence, jako např. flexibilní, snadno se učící atd. Vzhledem k rozdílům mezi profesními a osobními sociálními sítěmi je potřeba věnovat pozornost otázce, zda je možné usuzovat na osobnost nejen na základě osobních, ale také profesních sociálních sítí.

Ma a Leung (2018) realizovali výzkum na téma vliv osobnostních rysů a intenzity a způsobu užívání LinkedInu na získávání sociálního kapitálu. Jejich výzkumná otázka byla tedy opačná oproti většině výzkumů zmiňovaných v této práci. Neusuzovali na osobnost na základě způsobu užívání LinkedInu, ale usuzovali na způsob užívání LinkedInu na základě znalosti osobnostních rysů probandů. Konkrétně pracovali se souborem 301 uživatelů LinkedInu čínské národnosti, kteří vyplnili 20položkovou zkrácenou verzi Big Five dotazníku. Podle výsledků této studie lidé s vysokým skórem na škále přívětivosti, kteří zároveň často využívají LinkedIn za účelem zjišťování profesních informací od různých společností, vnímají, že získávají dobrý sociální kapitál. Lidé skórující vysoko na škále extravertze a/nebo neuroticismu používali LinkedIn častěji, a kromě sledování profesních informací ho používali i za účelem sebepropagace a budování strategických kontaktů.

Van de Ven, Bogaert, Serlie, Brandt a Denissen (2017) zkoumali přesnost lidských úsudků o osobnosti LinkedIn uživatelů vytvořených na základě jejich LinkedIn profilů. Konkrétně členové výzkumného týmu porovnávali odhady 10 hodnotitelů, studentů psychologie, ohledně osobnostních rysů dle modelu Big Five u 275 uživatelů LinkedInu, a tyto odhady porovnávali s výsledným skórem TIPI dotazníku, který uživatelé LinkedInu vyplnili. Sebraná data byla rozdělena na informace od pracujících a informace od studujících uživatelů LinkedInu. V rámci analýzy dat byla nalezena statisticky významná shoda mezi odhady a sebehodnocením pro rys extravertze v obou skupinách. Ve skupině pracujících byla kromě toho nalezena vysoká korelace pro rysy otevřenost vůči zkušenosti a přívětivost.

Rastogi a Nagappan (2016) se zaměřili na výzkum osobnosti uživatelů GitHubu. GitHub je profesní sociální síť, která umožňuje IT vývojářům sdílet jimi vytvořené kódy a spravovat projekty. Momentálně je na ní registrováno 31 miliónů uživatelů („GitHub: The world’s leading software development platform”, 2019). Rastogi a Nagappan (2016) analyzovali diskuze ohledně 243 projektů a dospěli k závěru, že vývojáři, jejichž aktivita byla ať už velmi vysoká nebo naopak velmi nízká, skórují výše na škále neuroticismu, oproti vývojářům, kteří přispívali do diskuze průměrně často. Dále autoři této studie zjistili, že vývojáři skórují tím výše na škálách svědomitosti a extraverze, čím déle jsou registrovaní na GitHubu. Pro škálu přívětivosti byl tento trend opačný.

Návrh výzkumného projektu

4 Výzkumný problém

Cílem navrhovaného výzkumu je ověřit možnosti odhadu osobnostních rysů jedince na základě jeho aktivity a podoby profilu na profesní sociální síti LinkedIn. Tato sociální síť byla jako zdroj informací vybrána proto, že poslední dobou vzniklo mnoho studií, které se zabývají odhadem osobnosti na základě osobních sociálních sítí (např. Segalin et al., 2017; Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011; Kern, Rickard, Seabrook, & Fulcher, 2018). Oproti tomu profesním sociálním sítím, tedy například LinkedInu, se takového výzkumného zájmu zatím nedostává. A to přesto, že je LinkedIn často používaný nástroj například v rámci nábory zaměstnanců. Podle Job Seeker National Study („Job Seeker National Study: Where Job Seekers Stand on the Economy, Job Security, and the Future of Work”, 2016) LinkedIn ke své práci v roce 2016 používalo 87 % recruiterů.

5 Design výzkumného projektu

Design navrhovaného výzkumu je do určité míry inspirovaný výzkumem, který realizovali Gosling, Augustin, Vazire, Holtzman a Gaddis (2011). Tito autoři zkoumali možnosti odhadu osobnostních rysů facebookových uživatelů dle modelu Big Five na základě jejich aktivity na Facebooku a podoby jejich facebookového profilu. Navrhovaný design se však od toho výše zmiňovaného liší v několika aspektech. Především bude vycházet z LinkedInu jakožto ze zdroje informací, nikoliv z Facebooku. Dále budou analyzovány informace z profilů v českém jazyce, nikoliv v anglickém jazyce. Za třetí bude k měření osobnosti využit jiný dotazník. Gosling, Augustine, Vazire, Holtzman a Gaddis (2011) použili TIPI, tedy zkrácenou verzi dotazníku Big Five, což je u podobného typu výzkumů častá volba (viz také van de Ven, Bogaert, Serlie, Brandt, & Denissen, 2017; Błachnio, Przepiorka, Senol-Durak, Durak, & Sherstyuk, 2017; Gezgin & Sen, 2012).

V rámci navrhovaného výzkumu však bude použit Hoganův osobnostní dotazník neboli HPI (z anglického Hogan Personal Inventory).

HPI vychází z modelu Big Five, ale dvě z původních pěti škál jsou rozděleny každá na dvě nové škály, takže se HPI celkově skládá ze sedmi škál. Konkrétně je extraverte rozdělena na sociabilitu a sebeprosazení a otevřenost vůči zkušenosti na zvědavost a učenlivost. Těchto sedm škál je dále možné dělit na 43 subškál, což poskytuje možnost detailnější interpretace výsledného skóru (Hogan, Hogan, & Warrenfeltz, 2007). Stručný popis škál HPI je následující:

Název škály HPI	Název škály Big Five	Interpretace vysokého skóru HPI	Interpretace nízkého skóru HPI
Stabilita	Neuroticismus (reverzně)	Jedinci s vysokým skórem zůstávají klidní a vyrovnaní ve stresových situacích.	Jedinci s nízkým skórem bývají citliví a vnímaví na zpětnou vazbu. Se stresem se mohou vyrovnávat obtížně.
Sebeprosazení	Extraverze	Jedinci s vysokým skórem se rádi ujímají vedoucích rolí, působí energicky a soutěživě.	Jedinci s nízkým skórem bývají tišší, trpěliví a ochotní přizpůsobit se ostatním.
Sociabilita	Extraverze	Jedinci s vysokým skórem bývají společenší, obvykle často a rádi navazují nové sociální vztahy.	Jedinci s nízkým skórem jsou méně hovorní, obvykle se nestávají středem pozornosti a nevdají jim pracovat o samotě.
Kooperativnost	Přívětivost	Jedinci s vysokým skórem jsou zaměřeni	Jedinci s nízkým skórem mohou mít

		na vztahy s ostatními, může pro ně být obtížné říkat „ne“.	tendenci jednat přímočaře a bez ohledu na ostatní.
Systematičnost	Svědomitost	Jedinci s vysokým skóreem bývají organizovaní, obvykle rádi plánují a chovají se podle pravidel.	Jedinci s nízkým skóreem bývají flexibilní, impulzivní a kreativní. Mohou se obtížně podřizovat pravidlům.
Zvídavost	Otevřenost vůči zkušenosti	Jedinci s vysokým skóreem bývají nápadití a bystří. Na druhou stranu mohou mít nižší odolnost vůči rutinní práci.	Jedinci s nízkým skóreem bývají realističtí a nezačnou se nudit ani při stereotypní činnosti.
Učelnivost	Otevřenost vůči zkušenosti	Jedinci s vysokým skóreem vyhledávají příležitosti k aktualizaci svých znalostí.	Jedinci s nízkým skóreem obvykle preferují tzv. „learning by doing“ oproti formálnímu vzdělávání.

Tabulka 1: Popis škál HPI (Hogan, Hogan, & Warrenfeltz, 2007).

Výhoda užití HPI místo obvykle používaného TIPI spočívá v podrobnější informační hodnotě díky většímu počtu škál a subškál při zachování možnosti porovnání výsledků s již provedenými studiemi, ve kterých byl použit TIPI či jiná verze dotazníku vycházejícího z modelu Big Five. HPI navíc na rozdíl od Big Five nepopisuje identitu jedince, tedy to, jaké má o sobě testovaný jedinec mínění, nýbrž jeho reputaci, tedy to, jak je testovaný jedinec pravděpodobně vnímaný svým okolím (Hogan, Hogan, & Warrenfeltz, 2007). HPI vychází z pracovního prostředí, což může být v kontextu výzkumu zaměřeného na profesní sociální síť jeho další výhodou. Dále je HPI dostupné i s českými normami. Nevýhoda užití HPI oproti

TUPI pak spočívá v času potřebném pro jeho vyplnění, neboť zabere přibližně 20 minut.

5.1 Metody získávání dat

Navrhovaný design výzkumu nevyžaduje seznání účastníků na jedno místo v jeden čas, což je jeho velká výhoda. Veškerou komunikaci tak bude možné realizovat online, konkrétně prostřednictvím zpráv na LinkedInu (vlastnictví LinkedIn profilu v českém jazyce je jednou z podmínek účasti na výzkumu). Zájemci o účast na výzkumu budou požádáni o odeslání LinkedIn žádosti ke spojení autorce výzkumu. Tento krok je nezbytný z toho důvodu, aby měla autorka výzkumu přístup k informacím uveřejněným na profilech účastníků výzkumu z pohledu tzv. spojení prvního stupně. Bez tohoto opatření by nebylo vzhledem k nastavení soukromí LinkedInu možné analyzovat informace z profilů účastníků výzkumu v plné šíři. Účastníci výzkumu budou upozorněni, že se jedná jen o dočasné opatření a že toto spojení mohou po skončení výzkumu ukončit.

Po navázání spojení na LinkedInu budou z profilů účastníků manuálně extrahovány a zaznamenávány do programu Excel následující charakteristiky a proměnné:

1. **Jméno, příjmení.** Tato informace je potřebná kvůli pozdějšímu párování s výsledky HPI a odpověďmi ze vstupního dotazníku.
2. **Počet LinkedIn spojení.** S kolika lidmi je účastník výzkumu propojen.
3. **Přítomnost profilové fotografie.**
 - ano x ne
4. **Profilová fotografie zobrazuje člověka.**
 - ano x ne
5. **Přítomnost fotografie na pozadí.**
 - ano x ne
6. **Upravená adresa URL.** Uživatelé LinkedInu si mohou upravit URL adresu svého profilu z automaticky vygenerované, která je obvykle poměrně dlouhá a obsahuje shluk písmen a číslic (např.

<https://www.linkedin.com/in/petrafouskova-a611a6139/>), na jiný formát, například <https://www.linkedin.com/in/petrafouskova/>.

○ ano x ne

7. **Podrobnost vyplnění.** Počet slov v sekcích souhrn, pracovní zkušenosti a vzdělání.
8. **Počet doporučení, která účastník výzkumu na síti LinkedIn udělil.**
9. **Počet doporučení, která účastník výzkumu obdržel.**
10. **Počet odborných článků, které účastník výzkumu napsal a sdílel za poslední rok.**
11. **Počet příspěvků, které účastník výzkumu sdílel za účelem sebepropagace za poslední rok.**
12. **Počet uvedených dovedností.**
13. **Počet uvedených zájmů.**

Následně účastníci obdrží LinkedIn zprávu s odkazem, který je přesměruje na online dotazník. V první části dotazníku budou účastníci vyzváni k poskytnutí informovaného souhlasu s účastí na výzkumu, jehož součástí bude i informace o zpracování osobních údajů. Informovaný souhlas bude účastníky zpravovat mimo jiné i o tom, jakým způsobem jim bude zaslána jejich výsledná zpráva k HPI, a že mají možnost využít nabídky 15minutové telefonické konzultace svého LinkedIn profilu (viz Příloha 2). Po poskytnutí informovaného souhlasu bude následovat vyplnění vstupního dotazníku, dotazujícího se účastníků na jejich pohlaví, věk, nejvyšší dosažené vzdělání a frekvenci a účel užívání LinkedInu (viz Příloha 3). V tomto dotazníku jsou pro kontrolu zahrnuty i věkové kategorie „méně než 15 let“ a „více než 55 let“. Data účastníků, kteří označí jednu z těchto dvou možností, však nebudou zahrnuta do zpracování dat. Dolní hranice 15 let je dána povoleným vstupem na pracovní trh v České republice. Horní hranice pak tím, že pouze malé množství uživatelů LinkedInu je starších 55 let. Podle průzkumu marketingové agentury Mojo Marketing tvořili uživatelé starší 55 let v roce 2017 7 % všech

uživatelů. Po vyplnění vstupního dotazníku budou účastníci výzkumu vyzváni k vyplnění HPI. Celý průběh sběru dat zobrazuje graf 1.



Graf 1. Průběh sběru dat

Po vyplnění HPI bude účastníkům zaslána prostřednictvím LinkedIn zpráv písemná narativní výsledná zpráva k HPI, ve které se budou moci dočíst interpretaci svých výsledků. Výsledné zprávy jsou generovány automaticky, tudíž jejich zaslání účastníkům výzkumu nebude časově náročné. Při odeslání výsledných zpráv bude účastníkům připomenuta možnost zažádání o konzultaci jejich LinkedIn profilu. Tato konzultace bude účastníkům výzkumu poskytována bezplatně autorkou výzkumu.

5.2 Hypotézy

Výzkum vychází z počtu proměnných, který představuje výzvu pro formulaci hypotéz tak, aby byly snadno srozumitelné. Rozhodla jsem se po vzoru Ma a Leunga (2018) uvádět pro větší přehlednost hypotézy souhrnně podle škály, ze které vycházejí. Domnívám se, že opačný způsob, tedy rozepsání sedmi níže uvedených hypotéz do mnoha dílčích hypotéz, by byl nepřehledný.

Výzkum Liu, Preoțiu-Pietra, Samani, Moghaddama a Ungara (2016) ukázal na pozitivní korelaci mezi vysokým skórem na škále neuroticismu a výběrem profilové fotografie na Twitteru, která nezobrazuje lidi, případně nezobrazuje jejich tváře. Očekávaná hypotéza pro škálu stability (reverzní neuroticismus v Big Five) je tedy následující:

H₁: **Stabilita** negativně souvisí s přítomností profilové fotografie, profilové fotografie, na které je přítomen člověk, a fotografie na pozadí.

Co se extraverze (tj. sebeprosazování a sociability v HPI) týče, Ma a Leung (2018) realizovali výzkum, který poukazuje na pozitivní korelaci mezi vysokým

skórem na této škále a frekvencí používání LinkedInu. Také podle výzkumu Goslinga, Augustina, Vazire, Holtzmana a Gaddiho (2011) používají sociální síť (v tomto případě Facebook, ale dá se předpokládat, že obdobné to bude i u jiných sociálních sítí) častěji extroverti než introverti. Extroverti měli dále větší počet přátel, častěji psali komentáře a aktualizovali informace o sobě. Jelikož škála sebeprosazování souvisí se soutěživostí a sebedůvěrou (Hogan, Hogan, & Warrenfeltz, 2007), je znění druhé hypotézy následující:

H₂: **Sebeprosazování** pozitivně koreluje s počtem obdržných doporučení, počtem uvedených dovedností a počtem příspěvků sdílených za účelem sebepropagace.

Vysoký skóre na škále sociability se pojí s frekvencí, s jakou lidé navazují nové kontakty a strategická spojení. Vysoce skórující jedinci bývají hovorní, přátelští a působí přístupným dojmem (Hogan, Hogan, & Warrenfeltz, 2007). Třetí hypotéza zní:

H₃: **Sociabilita** pozitivně koreluje s počtem hodin strávených týdně na LinkedInu, frekvencí používání LinkedInu, počtem spojení, počtem uvedených zájmů a podrobností vyplnění profilu (tzn. počtem slov v sekcích souhrn, pracovní zkušenosti a vzdělání).

Podle Augustina, Vazire, Holtzmana a Gaddiho (2011) koreluje pozitivně vysoký skóre na škále přívětivosti (kooperativnost v HPI) s frekvencí, s jakou uživatelé Facebooku navštěvují vlastní profily i profily ostatních lidí. Podle Hogana, Hoganové a Warrenfeltze (2007) se vysoký skóre na škále kooperativnosti pojí s pečováním o již navázané vztahy. Kooperativní lidé jsou společenější, působí příjemně a jsou obvykle týmoví hráči. Čtvrtá hypotéza tedy zní následovně:

H₄: **Kooperativnost** pozitivně koreluje s udílením doporučení ostatním uživatelům.

Výzkum Goslinga, Augustina, Vazire, Holtzmana a Gaddise (2011) našel pozitivní korelaci mezi svědomitostí (systematičností v HPI) a časem stráveným na různých facebookových stránkách. Důvodem může být to, že lidé s nízkým skórem

na škále svědomitosti na Facebooku prokrastinují. Liu, Preoțiu-Pietro, Samani, Moghaddam a Ungar (2016) zjistili, že mezi svědomitými lidmi je nejvyšší výskyt fotografií, na kterých se daná osoba usmívá a vypadá spokojeně. Komentují to tak, že vyjadřování pozitivních emocí sice není charakteristické pro tento osobnostní rys, ale naplňuje to očekávání, které lidé od profilových fotografií mohou mít. Vzhledem k tomu, se jsou lidé s vysokým skórem na škále systematickosti popisováni jako organizovaní, pečliví a plánující (Hogan, Hogan, & Warrenfeltze, 2007), vypadá pátá hypotéza následovně:

H₅: **Systematickosti** pozitivně koreluje s vyplněním kontaktních informací a upravením URL adresy profilu.

Oproti tomu lidé s vysokým skórem na škále otevřenosti vůči zkušenosti (zvědavost a učelnivost v HPI) mají podle výzkumu Liu, Preoțiu-Pietra, Samani, Moghaddama a Ungara (2016) tendenci volit profilové fotografie s absencí tváře. Výzkum Goslinga, Augustina, Vazire, Holtzmana a Gaddise (2011) (2011) ukázal na pozitivní korelace mezi otevřeností vůči zkušenosti a častým přidáváním a nahrazováním fotografií. Podle Hogana, Hoganové a Warrenfeltze (2007) se vysoký skór na škále zvědavost pojí s širokou paletou zájmů a vyhledáváním nových podnětů. Šestá hypotéza proto vypadá následovně:

H₆: **Zvědavost** pozitivně koreluje s počtem uvedených zájmů.

Učelniví lidé vyhledávají nové poznatky a odborné informace (Hogan, Hogan, & Warrenfeltz, 2007). Sedmá hypotéza proto zní:

H₇: **Učelnivost** pozitivně koreluje s počtem sdílených odborných článků.

5.3 Výzkumný soubor

Výzkumu se zúčastní 300 probandů. Ti budou pro zvýšení reprezentativnosti vybráni na základě kvótního výběru reprezentujícího Českou republiku pro kritéria pohlaví, věk a nejvyšší dosažené vzdělání. Zastoupení jednotlivých kategorií uvedené v tabulce níže je odvozeno z informací Českého statistického úřadu pro rok 2017 („Věk a vzdělání populace“, 2018). Účast na výzkumu bude dále podmíněna vlastnictvím LinkedIn profilu v českém jazyce.

Muži n = 150			
	Základní vzdělání a bez vzdělání	Střední vzdělání bez maturity, střední vzdělání s maturitou	Vysokoškolské vzdělání
15–24	12	12	3
25–34	3	27	12
35–54	3	60	18

Tabulka 2: Struktura obyvatelstva (muži). (Upraveno dle „Věk a vzdělání populace“, 2018).

Ženy n = 150			
	Základní vzdělání a bez vzdělání	Střední vzdělání bez maturity, střední vzdělání s maturitou	Vysokoškolské vzdělání
15–24	12	12	3
25–34	3	21	15
35–54	3	60	21

Tabulka 3: Struktura obyvatelstva (ženy). (Upraveno dle „Věk a vzdělání populace“, 2018).

Za účelem propagace projektu a oslovení dostatečného množství probandů bude vytvořena skupina na Facebooku a LinkedInu, ve které budou popsány podmínky účasti ve výzkumu. Dále budou vytvořeny propagační letáky, které budou šířeny v papírové i online podobě. V papírové podobě budou rozvěšeny na nástěnkách škol, úřadů práce, v personálních agenturách, ve výzkumných laboratořích Labels a Pless atd. Online pak bude plakát s nabídkou účasti na výzkumu distribuován prostřednictvím nejružnějších facebookových skupin (např. Respondenti, Praha – PRÁCE, BRIGÁDY, Přátelé filosofie a další – podle potřeby naplnění jednotlivých kvót). Výstupy této studie budou tedy zobecnitelné na uživatele LinkedInu ve věku 15–54 let, kteří mají profil v českém jazyce. Je ale potřeba brát v potaz určitá omezení v podobě relativně malého výzkumného souboru.

5.4 Metody zpracování a analýzy dat

Analýza dat bude provedena ve dvou krocích. V rámci první fáze budou ověřeny dílčí vztahy a hypotézy. Cílem této části je tedy testování hypotéz uvedených v předchozí kapitole, jinými slovy je cílem zjistit, zda existuje vztah mezi jednotlivými proměnnými, které se vztahují k aktivitě na LinkedInu a osobnostními rysy měřenými HPI. Tohoto cíle bude dosaženo korelační analýzou v programu SPSS, do kterého bude nahrána datová matice obsahující data ze vstupního dotazníku, výsledné skóry HPI a extrahované proměnné z LinkedInu. Korelační analýzu pro testování hypotéz použili ve svém výzkumu také Ma a Leung (2018) a Tadesse, Lin, Xu a Yang (2018).

Druhá fáze bude zaměřená na hledání komplexních predikčních modelů. Konkrétně bude zjištěna vysvětlená variance u jednotlivých škál HPI pomocí všech proměnných o aktivitě na LinkedInu, jejichž vztah s danou škálou byl hypotetizován výše, a to mnohonásobnou lineární regresí. Za tímto účelem bude vytvořeno sedm modelů, každý pro jednu škálu HPI. Lineární regresi zvolili pro své predikční modely také Ma a Leung (2018); Liu, Preoțiu-Pietro, Samani, Moghaddam a Ungar (2016) a Kosinski, Stillwell a Graepel (2013). Pro prevenci multikolinearity bude nejprve provedena explorační faktorová analýza. Pokud budou paralelní analýzou a vizuálním srovnáním sutinového grafu nalezeny latentní faktory sdružující vícero proměnných, budou tyto faktory použity v dalších analýzách místo samotných proměnných. Ověření modelu může být provedeno například porovnáním predikovaných a skutečných HPI skóru užitím Pearsonova korelačního koeficientu (stejný postup volili také Liu, Preoțiu-Pietro, Samani, Moghaddam, & Ungar, 2016), nebo pomocí RSME (toto také navrhuje Liu, Preoțiu-Pietro, Samani, Moghaddam, & Ungar, 2016). Regresní modely nejlépe odpovídající datům budou nalezeny pomocí krokového (stepwise) algoritmu výběru prediktorů.

6 Etika výzkumu

Výzkum bude realizován v souladu s Etickým kodexem psychologické profese, vydaným Českomoravskou psychologickou společností („Etický kodex psychologické profese: Verze 2017“, 2017). Výzkumu se budou účastnit pouze účastníci, kteří k tomu poskytnou písemný informovaný souhlas s účastí na výzkumu, ve kterém je zahrnuta informace o zpracování osobních údajů v souladu s obecným nařízením o ochraně osobních údajů č. 2016/679 (GDPR). Informovaný souhlas je uvedený v příloze 2. Důvodem pro využití informace o zpracování osobních údajů je fakt, že v rámci výzkumu budou analyzovány LinkedIn profily probandů obsahující jejich osobní údaje (jméno, příjmení, dosažené vzdělání a pracovní zkušenosti, současná pracovní pozice, praxe a brigády, kontaktní informace a další související položky). V rámci výzkumu budou zpracovány jen poskytnuté informace – tedy výsledky osobnostní psychodiagnostiky, vstupního dotazníku a poskytnutý obsah LinkedIn profilu. Získané osobní údaje nebudou poskytovány třetím stranám a budou po skončení výzkumu anonymizovány.

Účastníci obdrží výsledky absolvované psychodiagnostiky ve formě písemné narativní výsledné zprávy. Tato výsledná zpráva je generována automaticky počítačem na základě vyplněného dotazníku. Kromě toho bude účastníkům nabídnuta 15minutová konzultace LinkedIn profilu, která bude probíhat telefonicky po skončení sběru dat. Po ukončení sběru dat a zpracování výsledků bude také pro účastníky zorganizováno informační setkání, v jehož rámci se budou moci v případě zájmu dozvědět více o výzkumu, jehož se zúčastnili.

7 Diskuse

Výzkumy zaměřující se na možnosti odhadu osobnostních rysů a inteligence na základě digitálních stop a výsledků činnosti se poslední dobou dostávají do středu vědeckého zájmu. Například Gosling, Augustine, Vazire, Holtzman a Gaddi (2011) zjišťovali vztah mezi aktivitou uživatelů na Facebooku a jejich osobnostními rysy. Liu, Preoțiu-Pietro, Samani, Moghaddam a Ungar (2016) zkoumali možnosti odhadu osobnostních rysů uživatelů Twitteru na základě

jejich profilových fotografií. Do třetice, Ma a Leung (2018) se zaměřili na možnosti odhadu osobnostních rysů uživatelů LinkedInu na základě jejich aktivity na této síti.

Navzdory výše zmíněnému by proponovaný projekt mohl obohatit dosavadní vědecké poznání hned v několika oblastech. Jednak vychází z LinkedIn profilů v českém jazyce, tudíž jsou jeho výsledky zobecnitelné na českou populaci. Samozřejmě jen do jisté míry, neboť studie neaspíruje na získání 100 % reprezentativního výběrového souboru. Vzhledem k tomu, že výzva k účasti na výzkumu bude šířená formou letáků, nebudou mít všichni lidé stejnou možnost se o této výzvě dozvědět a být tedy zahrnuti do výběrového souboru. Do výzkumu nebudou dále zahrnuti lidé, kteří nemají LinkedIn profil. Toto všechno mluví v neprospěch reprezentativnosti výběrového souboru. Na druhou stranu, reprezentativnost je podpořena kvótním výběrem v kritériích věk, pohlaví a nejvyšší dosažené vzdělání. Pro srovnání, Gosling, Augustine, Vazire, Holtzman a Gaddi (2011) měli probandy studenty. Ma a Leung (2018) použili pro výběr probandů metodu sněhové koule (snowball sampling). Kučera et al. (2018) v rámci projektu CPACT sice také zvolili kvótní výběr, předmětem jejich analýzy však nebyl obsah sociálních sítí, nýbrž texty vzniklé uměle v psychologické laboratoři, u kterých je tedy zpochybnitelná ekologická validita.

Výhodou navrhovaného výzkumu je také možnost porovnání výsledků se zahraničními studii, které vychází z pětifaktorového modelu osobnosti, neboť proponovaný projekt předpokládá využití Hoganova osobnostního dotazník, který je také založen na modelu Big Five. Rozšíření původního pětifaktorového modelu na sedmifaktorový v kombinaci s informacemi ze 43 subškál však nabízí oproti jiným výzkumům podrobnější informační hodnotu. Hoganův osobnostní dotazník sestává ze 168 položek (Hogan, Hogan, & Warrenfeltz, 2007), což klade větší časové nároky na probandy v porovnání s 10položkovým TIPI. Zároveň to ale podporuje psychometrické vlastnosti této metody. Také je dostupný i s českými normami. Na druhou stranu, zvolená psychodiagnostika by pravděpodobně patřila mezi největší nákladové položky ve výzkumném rozpočtu a musela by být financována z grantů, případně by bylo třeba jednat o získání volných přístupových údajů za účelem výzkumu s distributory této metody.

Co se metody zpracování a analýzy dat týče, je možné spatřovat jejich limit v tom, že návrh předpokládá analýzu dat pomocí lineární regrese. Ač tuto metodu volili i jiní výzkumníci (Ma & Leung, 2018; Liu, Preoțiu-Pietro, Samani, Moghaddam, & Ungar, 2016; Kosinski, Stillwell, & Graepel, 2013), jedná se o základní statistickou metodu, která je v silách studenta bakalářské psychologie. Přesnější predikce by bylo patrně možné dosáhnout sestavením strukturálního modelu, což ve svém výzkumu zvolili například Chen, Widjaja a Yen (2015). K tomu bych ovšem v současné chvíli potřebovala absolvovat doplňkové kurzy statistiky, případně projekt intenzivně konzultovat s někým zkušenějším.

Mezi limity projektu patří také velikost výzkumného souboru, který čítá 300 respondentů. V kontextu analýzy digitálních stop a výsledků činnosti vznikly studie s výrazně početnějšími soubory dat, jmenovitě například projekt MyPersonality s daty od 6 000 000 uživatelů Facebooku („MyPersonality.org“), či výzkum od Liu, Preoțiu-Pietra, Samani, Moghaddama a Ungara (2016) s daty 66 000 uživatelů Twitteru.

I přesto, že studie předpokládá výběrový soubor o velikosti $n = 300$, je třeba počítat s náročností spojenou se zaplněním všech předem stanovených kvót vzhledem ke zvolené metodologii výběru probandů. Dalším omezením navrhovaného výzkumu je také horní věková hranice respondentů, stanovená na 55 let. Pokud by na navrhovaný výzkum navázala další studie, zjišťující možnosti odhadu osobnostních rysů uživatelů LinkedInu starších 55 let, bylo by dozajista zajímavé porovnat výsledky obou studií.

Další omezení spočívá v tom, že seznam proměnných, jejichž schopnost predikce osobnostních rysů bude v rámci výzkumu testována, není vyčerpávající. Při podrobnějším zkoumání by byly patrně identifikovány další proměnné, které by mohly zpřesnit výpovědní hodnotu predikčního modelu. Počet sledovaných proměnných je však porovnatelný s jinými studiemi, například Ma a Leung (2018) si zvolili 13 proměnných.

Další úskalí navrhovaného výzkumu spočívá v možném ovlivnění podoby dat ze strany účastníků. Účastníci výzkumu, vědomí si skutečnosti, že jejich profil bude podroben bližší nespecifikovanému zkoumání, mohou mít tendenci svůj profil

kvůli tomuto faktu upravit (například aktualizovat informace, změnit profilovou fotografii atd). To by mělo samozřejmě negativní vliv na ekologickou validitu studie. Gosling, Augustine, Vazire, Holtzman a Gaddis (2011) předešli tomuto efektu tím způsobem, že si dopředu stáhli informace alespoň z veřejných profilů účastníků výzkumu, a teprve poté účastníky informovali, že výzkum zahrnuje i informace z jejich profilů.

Dále je třeba zmínit, že výzkum počítá se zpracováním osobních údajů. Do výzkumu jsou také zahrnuty informace z absolvované psychodiagnostiky. Jedná se tedy o citlivé informace a je proto na místě si přiznat, že navrhovaný souhlas s účastí na výzkumu, který obsahuje i informaci se zpracováním osobních údajů, by měl být před distribucí účastníkům zkontrolován a případně upraven osobou s relevantním právnickým vzděláním. Krom toho, zvýšená citlivost veřejnosti v oblasti ochrany osobních údajů, způsobená tím, že před rokem vstoupilo v účinnost Nařízení Evropského parlamentu a Rady (EU) 2016/679 (GDPR), by mohla negativně ovlivnit ochotu probandů účastnit se výzkumu, který není anonymní. V praxi by pak bylo potřeba citlivě komunikovat, jaká data a za jakým účelem budou zpracovávána, a počítat s tím, že to může některé probandy odradit.

Pokud by navrhovaná studie našla vztahy mezi zkoumanými proměnnými, mohly by být tyto poznatky uplatněny v rámci psychologie práce a organizace, například za účelem zefektivnění náborových a výběrových procesů. Analýza LinkedIn profilů uchazečů by mohla být například využita v kontextu předvýběru vhodných kandidátů. Alternativně by výstupy navrhované studie mohly sloužit jako výchozí bod pro další vědecké zkoumání této problematiky. Tedy i v případě, pokud nebudou nalezeny vztahy mezi zkoumanými proměnnými, bude to z vědeckého hlediska přínosné zjištění. Co se literárně přehledové části týče, ta by mohla být v rámci práce s větším rozsahem (například diplomové práce) rozšířena o další trendy na poli psychometrie, ke kterým patří například game-based assessment, či adaptivní testování založené na IRT (item response theory – teorie odpovědi na položku).

Na tento výzkum je možné navázat studii zaměřenými na porovnání odhadu osobnostních vlastností z různých sociálních sítí, případně se zaměřit na

lokální odlišnosti a porovnat možnosti odhadu osobnostních rysů z LinkedIn profilů v různých jazycích. Další možností je podrobit informace získané z LinkedIn profilů počítačové analýze textu, například prostřednictvím programu LIWC. Stejně tak by mohlo být zajímavé extrahovat detaily z LinkedIn profilových fotografií a fotografií na pozadí (například programem Face++) a podrobit je analýze podobně, jako to ve svém výzkumu učinili Segalin et al. (2017).

Závěr

Cílem předkládaného textu bylo na základě odborné literatury, článků a studií předložit ucelený pohled na možnosti využití digitálních stop k odhadu osobnostních vlastností a inteligence. Tohoto cíle bylo dosaženo v literárně přehledové části této práce. V rámci prvních tří kapitol byly popsány možnosti analýzy textu, včetně výhod a omezení plynoucích z této metody. Také zde byly shrnuty významné studie zabývající se odhadem osobnostních rysů a inteligence na základě fotografií. Práce obsahuje podrobnější informace o výzkumech zaměřených na odhad osobnosti prostřednictvím analýzy chování uživatelů osobních a profesních sociálních sítí, jako jsou například Facebook, Twitter či LinkedIn.

Proponovaný projekt je zaměřen na odhad osobnostních vlastností v sedmifaktorovém pojetí dle HPI na základě aktivity uživatelů na LinkedInu. Předpokládá účast 300 účastníků výzkumu, kteří mají LinkedIn profil v českém jazyce. Z profilů účastníků budou manuálně extrahovány proměnné, které budou spolu s výstupy ze vstupního a Hoganova osobnostního dotazníku analyzovány v programu SPSS. Analýza dat bude kromě ověření dílčích vztahů a hypotéz zaměřená na hledání komplexních predikčních modelů.

Co se přínosu práce týče – vzhledem k tomu, že literárně přehledová část práce obsahuje shrnutí nejnovějších studií, převážně zahraničních, zabývajících se měřením osobnosti na základě analýzy digitálních stop a výsledků činnosti, může tato práce sloužit jako přehled čtenářům, kteří se zajímají o aktuální dění na poli moderní psychometrie. Konkrétně čtenáři získají vhled do možností a limitů analýzy textů, fotografií a uživatelského chování na internetu. Realizace výzkumu, navrženého v empirické části této práce, a s tím související vytvoření predikčního modelu k odhadu osobnostních rysů na základě profilu na LinkedInu, by mohla obohatit současný stav poznání a zefektivnit praxi v oblasti náboru a výběru zaměstnanců.

Seznam použité literatury

AnalyzeWords [Online]. Retrieved March 15, 2019, from <http://www.analyzewords.com/index.php/>.

APA (2010). *Publication Manual of the American Psychological Association* (6th Ed.). Washington, DC: American Psychological Association.

Back, M. D., Schmukle, S. C., & Egloff, B. (2008). How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses [Online]. *Journal Of Research In Personality*, 42(4), 1116–1122. <https://doi.org/10.1016/j.jrp.2008.02.001>.

Baddeley, J. I., Daniel, G. R., & Pennebaker, J. W. (2011). How Henry Hellyer's Use of Language Foretold His Suicide [Online]. *Crisis: The Journal Of Crisis Intervention And Suicide Prevention*, 32(5), 288–289. Retrieved from <https://sfx.is.cuni.cz/sfxlcl3?ID=doi:10.1027/0227-5910.a000092&genre=article&atitle=How%20Henry%20Hellyer%E2%80%99s%20Use%20of%20Language%20Foretold%20His%20Suicide&title=Crisis%3A%20The%20Journal%20of%20Crisis%20Intervention%20and%20Suicide%20Prevention&issn=02275910&isbn=&volume=32&issue=5&date=20110101&aulast=Jenna%20L.,%20Baddeley&page=288&pages=288-289&sid=EBSCO:Journals%40OVID:edsovi.00003243.201132050.00009&sv.c.fulltext=yes>.

Błachnio, A., Przepiorka, A., Senol-Durak, E., Durak, M., & Sherstyuk, L. (2017). The role of personality traits in Facebook and Internet addictions: A study on Polish, Turkish, and Ukrainian samples [Online]. *Computers In Human Behavior*, (68), 269–275. <https://doi.org/10.1016/j.chb.2016.11.037>.

Boyd, R. I., & Pennebaker, J. W. (2017). Language-based personality: a new approach to personality in a digital world [Online]. *Current Opinion In Behavioral Sciences*, (18), 63–68. Retrieved from https://ac-els-cdn-com.ezproxy.is.cuni.cz/S2352154617300487/1-s2.0-S2352154617300487-main.pdf?_tid=3b4d94f6-9c22-4b33-b162-0bab87d1b7e6&acdnat=1552731597_3fdfeaf8e083759971d5b2a38b6798b1.

Cantador, I., Fernández-Tobías, I., & Bellogín, A. (2013). Relating personality types with user preferences in multiple entertainment domains [Online]. *Ceur Workshop Proceedings*, (997). Retrieved from <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/>.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic Markers of Psychological Change Surrounding September 11, 2001 [Online]. *Psychological Science*, 15(10), 687–693. Retrieved from <http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=4&sid=c6f8aefb-580d-4dc2-a651-c3263b847d07%40sdc-v-sessmgr01>.

Etický kodex psychologické profese: Verze 2017 [Online]. (2017). Českomoravská psychologická společnost – Etická komise. Retrieved from <https://cmps.ecn.cz/EK/Eticky-kodex-psychologicke-profese-12-2017.pdf>.

Evans, D. C., Gosling, S. D., & Carroll, A. (2008). What Elements of an Online Social Networking Profile Predict Target- Rater Agreement in Personality Impressions? [Online]. Association for the Advancement of Artificial Intelligence. Retrieved from <https://pdfs.semanticscholar.org/5695/f2b2f5467017e34c5d97129f2572a0cce45.pdf>.

Farnadi, G., Sitaraman, G., Sushmita, S., Celli, F., Kosinski, M., Stillwell, D., et al. (2016). Computational personality recognition in social media [Online]. *User Modeling*, 26(2/3), 109–142. <https://doi.org/10.1007/s11257-016-9171-0>.

Garcia, D., & Sikström, S. (2016). Friend or Worker? Descriptions of One's Personality in LinkedIn [Online]. *Personality And Individual Differences*, 101, 480–480. Retrieved from <https://doi.org/10.1016/j.paid.2016.05.144>.

Gezgin, U. başar, & Sen, N. yaw. (2012). "Keep In E-Touch": Personality and Facebook Use [Online]. *Akdeniz Iletisim*, (18), 123–139. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=8&sid=b08b2ad1-87c2-4f0f-862f-4cf30af1c32b%40sessionmgr104>.

Gill, A. J., & Oberlander, J. (2003). Perception of e-mail personality at zero-acquaintance: Extraversion takes care of itself; Neuroticism is a worry [Online]. *Proceedings Of The 25Th Annual Meeting Of The Cognitive Science Society*, 2003, 456–461. Retrieved from <http://csjarchive.cogsci.rpi.edu/proceedings/2003/pdfs/100.pdf>.

GitHub: The world's leading software development platform [Online]. (2019). Retrieved April 19, 2019. Retrieved from <https://github.com/>.

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of Personality in Online Social Networks: Self-Reported

Facebook-Related Behaviors and Observable Profile Information [Online]. *Cyberpsychology, Behavior And Social Networking*, 14(9), 483–488. Retrieved from <https://doi.org/10.1089/cyber.2010.0087>.

Groom, C. J., & Pennebaker, J. W. (2005). The Language of Love: Sex, Sexual Orientation, and Language Use in Online Personal Advertisements [Online]. *Sex Roles*, 52(7/8), 447–461. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=62f52a43-89e0-4dc6-a3c4-9ea221b32219%40pdc-v-sessmgr02>.

Hogan, R., Hogan, J., & Warrenfeltz, R. (2007). *The Hogan Guide: Interpretation and Use of Hogan Inventories*. Hogan Assessment Systems.

Chen, J., Widjaja, A., & Yen, D. (2015). Need for Affiliation, Need for Popularity, Self-Esteem, and the Moderating Effect of Big Five Personality Traits Affecting Individuals' Self-Disclosure on Facebook [Online]. *International Journal Of Human-Computer Interaction*, 31(11), 815–831. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=8aa16e60-372c-43a5-8984-cbfc1d15c965%40sessionmgr104>.

Ikeda, K., Hattori, G., Ono, C., Asoh, H., & Higashino, T. (2013). Twitter user profiling based on text and community mining for market analysis [Online]. *Knowledge-Based Systems*, 51, 35–47. Retrieved from <https://www.sciencedirect.com.ezproxy.is.cuni.cz/science/article/pii/S0950705113002025>.

Internet Live Stats [Online]. Retrieved March 10, 2019, from <http://www.internetlivestats.com/>.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., & Pennebaker, J. W. (2011). Language Style Matching Predicts Relationship Initiation and Stability [Online]. *Psychological Science*, 22(1), 39–44. Retrieved from https://www-jstor-org.ezproxy.is.cuni.cz/stable/40984604?read-now=1&seq=1#page_scan_tab_contents.

James W. Pennebaker [Online]. (c1996–2018). Retrieved March 15, 2019, from <https://pennebaker.socialpsychology.org/>.

Job Seeker National Study: Where Job Seekers Stand on the Economy, Job Security, and the Future of Work [Online]. (2016). Jobvite. Retrieved from https://www.jobvite.com/wp-content/uploads/2016/03/Jobvite_Jobseeker_Nation_2016.pdf.

Kaurin, A., Heil, L., Wessa, M., Egloff, B., & Hirschmueller, S. (2018). Selfies reflect actual personality – Just like photos or short videos in standardized lab conditions [Online]. *Journal Of Research In Personality*, 76, 154–164. Retrieved from <https://www-sciencedirect-com.ezproxy.is.cuni.cz/science/article/pii/S0092656618302484>.

Kern, M., Rickard, N. S., Seabrook, E. M., & Fulcher, B. D. (2018). Predicting depression from language-based emotion dynamics: Longitudinal analysis of facebook and twitter status updates [Online]. *Journal Of Medical Internet Research*, 20(5). <https://doi.org/10.2196/jmir.9267>.

Kleisner, K., Chvátalová, V., & Flegr, J. (2014). Perceived intelligence is associated with measured intelligence in men but not women [Online]. *Plos One*, 9(3), 1–7. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=4&sid=502d64a5-1622-4292-b3e3-59573b1e39ed%40pdc-v-sessmgr01>.

Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks [Online]. *Machine Learning*, 95(3), 357–380. Retrieved from <https://link-springer-com.ezproxy.is.cuni.cz/article/10.1007/s10994-013-5415-y>.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior [Online]. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 110(15), 5802–5805. Retrieved from https://www-jstor-org.ezproxy.is.cuni.cz/stable/42590309?read-now=1&seq=1#page_scan_tab_contents.

Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes [Online]. *Psychological Methods*, 21(4), 493–506. Retrieved from <http://eds.a.ebscohost.com/eds/detail/detail?vid=3&sid=bbe10612-6208-4ec8-8dd8-97c23364f064%40sdc-v-sessmgr02&bdata=Jmxhbm9Y3Mmc2l0ZT1lZHMtbG12ZSZzY29wZT1zaXRl#AN=2016-57141-003&db=pdh>.

Kučera, D., Havigerová, J. M., Haviger, J., Cvrček, V., Komrsková, Z., Lukeš, D., et al. (2018). *Výzkum CPACT: počítačová psycholinguistická analýza českého textu*. České Budějovice: Pedagogická fakulta Jihočeská univerzity.

Lambiotte, R., & Kosinski, M. (2014). Tracking the Digital Footprints of Personality [Online]. *Proceedings Of The Ieee*, 102(12), 1934–1939. Retrieved from <https://ieeexplore-ieee-org.ezproxy.is.cuni.cz/document/6939627/>.

MyPersonality.org [Online]. Retrieved March 11, 2019, from <https://sites.google.com/michalkosinski.com/mypersonality>.

LinkedIn Demographics [Online]. (2017). *Marketing Mojo*. Retrieved from https://www.marketing-mojo.com/wp-content/uploads/2017/04/Marketing-Mojo_LinkedIn_Demographics_Infographic_2017.pdf.

Liu, L., Preotjiuc-Pietro, D., Samani, Z. R., Moghaddam, M., & Ungar, L. (2016). Analyzing Personality through Social Media Profile Picture Choice [Online]. In *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*. Pennsylvania: Association for the Advancement of Artificial Intelligence. Retrieved from <https://wwbp.org/papers/persimages16icwsm.pdf>.

LIWC [Online]. Retrieved March 15, 2019, from <http://liwc.wpengine.com/>.

Ma, S. Q., & Leung, L. (2018). The Impacts of Personality Traits, Use Intensity and Features Use of LinkedIn on Bridging Social Capital [Online]. *Applied Research In Quality Of Life*, 1–20. <https://doi.org/10.1007/s11482-018-9635-y>.

Pelled, A., Zilberstein, T., Tsirulnikov, A., Pick, E., Patkin, Y., & Tal-or, N. (2017). Textual Primacy Online: Impression Formation Based on Textual and Visual Cues in Facebook Profiles [Online]. *American Behavioral Scientist*, 61(7), 672–687. Retrieved from <https://journals-sagepub-com.ezproxy.is.cuni.cz/doi/full/10.1177/0002764217717563>.

Pennebaker, J. W. (2011a). *The Secret Life of Pronouns: What Our Words Say About Us*. New York: Bloomsbury Press.

Pennebaker, J. W. (2011b). The Secret Life of Pronouns [Online]. *New Scientist*, 211(2828), 42–45. Retrieved from [https://doi.org/10.1016/S0262-4079\(11\)62167-2](https://doi.org/10.1016/S0262-4079(11)62167-2).

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays [Online]. *Plos One*, 9(12), 1–10. Retrieved from <http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=faf70df7-d077-463a-a672-a49367e31a24%40sdc-v-sessmgr04>.

Pennebaker, J. W., & King, L. A. (1999). Linguistic Styles: Language Use as an Individual Difference [Online]. *Journal Of Personality And Social Psychology*, 77(6), 1296–1297. Retrieved from <http://web.b.ebscohost.com.ezproxy.is.cuni.cz/ehost/pdfviewer/pdfviewer?vid=1&sid=4899e450-aa7e-435e-a295-cfd9056664bf%40pdc-v-sessmgr01>.

Qiu, L., Lu, J., Yang, S., Ramsay, J., Qu, W., & Zhu, T. (2017). Personality expression in Chinese language use [Online]. *International Journal Of Psychology*, 52(6), 463–472. <https://doi.org/10.1002/ijop.12259>.

Rastogi, A., & Nagappan, N. (2016). On the Personality Traits of GitHub Contributors [Online]. In *27th International Symposium on Software Reliability Engineering (ISSRE)* (pp. 77–86). Ottawa, (Canada): IEEE. Retrieved from <https://ieeexplore-ieee-org.ezproxy.is.cuni.cz/xpl/mostRecentIssue.jsp?punumber=7774240>.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L. D., Ramones, S. M., Agrawal, M., et al. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach [Online]. *Plos One*, 8(9), 1–16. <https://doi.org/10.1371/journal.pone.0073791>.

Segalin, C., Celli, F., Polonio, L., Kosinski, M., Stillwell, D., Sebe, N., et al. (2017). What your Facebook Profile Picture Reveals about your Personality [Online]. *Computer Science*. Retrieved from <https://arxiv.org/pdf/1708.01292.pdf>.

Stirman, S. W., & Pennebaker, J. W. (2001). Word Use in the Poetry of Suicidal and Nonsuicidal Poets [Online]. *Psychosomatic Medicine*, 63(4), 517–522.

Tadesse, M., Lin, H., Xu, B., & Yang, L. (2018). Personality Predictions Based on User Behavior on the Facebook Social Media Platform [Online]. *Ieee Access*, 6, 61959–61969. Retrieved from <https://ieeexplore-ieee-org.ezproxy.is.cuni.cz/document/8494744>.

Talamas, S. N., Mavor, K. I., & Perrett, D. I. (2016). Blinded by beauty: Attractiveness bias and accurate perceptions of academic performance [Online]. *Plos One*, 11(2), 1–18. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=7&sid=502d64a5-1622-4292-b3e3-59573b1e39ed%40pdc-v-sessmgr01>.

Talamas, S. N., Mavor, K. I., Perrett, D. I., Sundelin, T., & Axelsson, J. (2016). Eyelid-Openness and mouth curvature influence perceived intelligence beyond

attractiveness [Online]. *Journal Of Experimental Psychology*, 145(5), 1–18. Retrieved from <http://eds.b.ebscohost.com/eds/pdfviewer/pdfviewer?vid=9&sid=502d64a5-1622-4292-b3e3-59573b1e39ed%40pdc-v-sessmgr01>.

Utz, S. (2010). Show me your friends and I will tell you what type of person you are: How one's profile, number of friends, and type of friends influence impression formation on social network sites [Online]. *Journal Of Computer-Mediated Communication*, 15(2), 314–335. Retrieved from <https://academic.oup.com/jcmc/article/15/2/314/4064817>.

Van de Ven, N., Bogaert, A., Serlie, A., Brandt, M., & Denissen, J. (2017). Personality perception based on LinkedIn profiles [Online]. *Journal Of Managerial Psychology*, 32(6), 418–430. Retrieved from <https://www.emeraldinsight.com/doi/pdfplus/10.1108/JMP-07-2016-0220>.

Vazire, S., & Gosling, S. D. (2004). E-Perceptions: Personality Impressions Based on Personal Websites [Online]. *Journal Of Personality And Social Psychology*, 87(1), 123–132. Retrieved from <http://web.a.ebscohost.com.ezproxy.is.cuni.cz/ehost/pdfviewer/pdfviewer?vid=1&sid=05894d7e-50d6-4159-b669-4f90cacfc0a8%40sdc-v-sessmgr04>.

Věk a vzdělání populace [Online]. (2018). Český statistický úřad. Retrieved from <https://www.czso.cz/documents/10180/61565960/25013018001.pdf/ca1a4541-edf2-400e-ac91-f1f58545a290?version=1.0>.

Wei, X., & Stillwell, D. (2016). How smart does your profile image look? Estimating intelligence from social network profile images [Online]. <https://doi.org/10.1145/3018661.3018663>.

Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). Learning From Text: Matching Readers and Texts by Latent Semantic Analysis [Online]. *Discourse Processes*, 25(2), 309–336. <https://doi.org/10.1080/01638539809545030>.

Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers [Online]. *Journal Of Research In Personality*, 44(3), 363–373. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2885844/>.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans [Online]. *Pnas*, *112*(4), 1036–1040. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4313801/>.

Seznam tabulek

Tabulka 1: Popis škál HPI	31
Tabulka 2: Struktura obyvatelstva (muži)	37
Tabulka 3: Struktura obyvatelstva (ženy).....	37

Seznam použitých zkratk

CPACT: Computational Psycholinguistic Analysis of Czech Text (komputační psycholingvistická analýza českého textu)

HPI: Hogan Personal Inventory (Hoganův osobnostní dotazník)

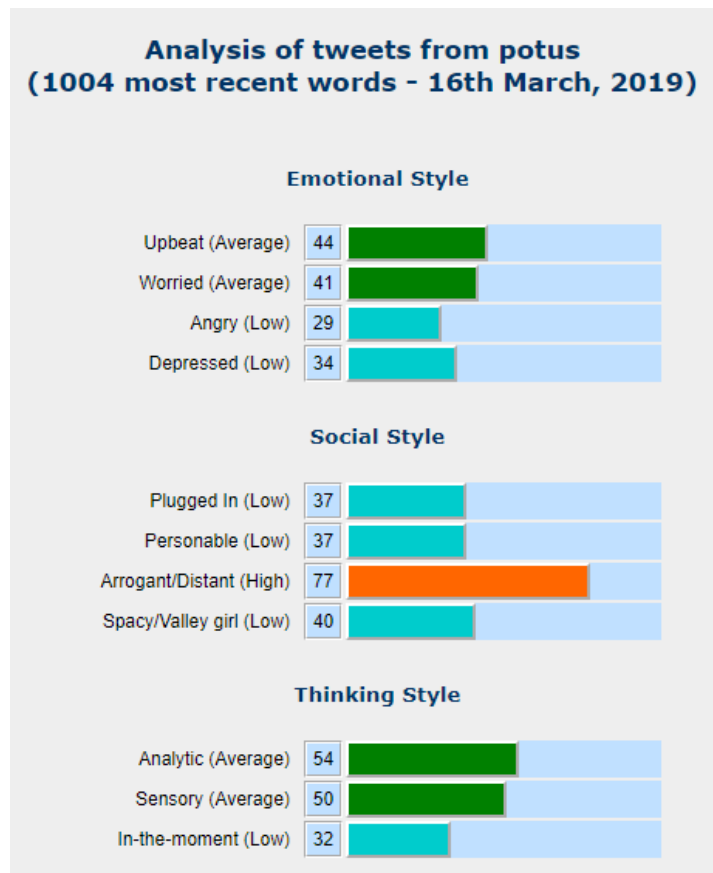
IPIP: International Personality Item Pool

LIWC: Linguistic Inquiry and Word Count

LSM: Language Style Matching

TIPI: Ten Item Personality Inventory

Příloha 1



Obrázek 1: Analýza twitterového profilu Donalda Trumpa („AnalyzeWords“)

Příloha 2

Informovaný souhlas s účastí na výzkumu

Dobrý den,

vítáme Vás na stránkách výzkumu zaměřeném na používání LinkedInu. Děkujeme za Váš zájem účastnit se výzkumu, velmi si toho vážíme.

Výzkum sestává ze dvou částí. Nejprve budete požádáni o vyplnění vstupního dotazníku, což by mělo zabrat zhruba 5 minut. Následně Vám budou poskytnuty přístupové údaje k vyplnění Hoganova osobnostního dotazníku, jehož vyplnění Vám zabere zhruba 20 minut. Po vyplnění obou dotazníků Vám do zpráv zašleme výslednou zprávu, ve které se budete moci seznámit s Vašimi výsledky z Hoganova osobnostního dotazníku. Krom toho si Vám dovoluujeme jako poděkování za Váš čas nabídnout zdarma 15minutovou telefonickou konzultaci Vašeho LinkedIn profilu. Pokud máte o tuto konzultaci zájem, dejte nám vědět prostřednictvím zpráv na LinkedInu Vaše časové preference.

Dále si Vás dovoluujeme v souladu s Nařízením Evropského parlamentu a Rady (EU) 2016/679 informovat, že v souvislosti s Vaší účastí na výzkumu dojde ke zpracování Vašich osobních údajů, a to v rozsahu, v jakém jsou uvedené na Vašem LinkedIn profilu. V případě jakýchkoliv dotazů se prosím obraťte na výzkumný tým.

Účast na tomto výzkumu je dobrovolná a můžete ji kdykoliv ukončit.

Pokračováním na další stranu udělujete souhlas s účastí na tomto výzkumu.

Děkujeme za Váš čas,

Výzkumný tým

Příloha 3

VSTUPNÍ DOTAZNÍK

Odpovězte prosím na následující otázky týkající se Vaší osoby.

1. Jméno a příjmení:

-

2. Pohlaví:

- žena
- muž

3. Věk

- méně než 15 let
- 15–24 let
- 25–34 let
- 35–54 let
- více než 54 let

4. Nejvyšší dosažené vzdělání:

- základní vzdělání
- střední odborné vzdělání
- střední všeobecné vzdělání
- vysokoškolské vzdělání
- jiné – doplňte

5. Jak často používáte LinkedIn?

- denně
- několikrát týdně
- 1krát týdně
- 1krát za měsíc

- méně než 1krát za měsíc

6. Kolik hodin týdně strávíte na LinkedInu?

- jednu hodinu a méně
- 2–3 hodiny
- 4–9 hodin
- více než 10 hodin

7. Za jakým účelem používáte LinkedIn? Označte všechny správné odpovědi.

- hledání zaměstnání
- hledání obchodních příležitostí
- headhunting
- výzkum
- nástroj posílení osobní značky
- jiný – doplňte