

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

MAGISTRSKO DELO

**ANALIZA STANJA IN POSLOVNIH PRILOŽNOSTI UPORABE
MASOVNIH PODATKOV V VELIKIH SLOVENSKIH PODJETJIH**

Ljubljana, september 2018

MONIKA ČURIN

IZJAVA O AVTORSTVU

Podpisana Monika Čurin, študentka Ekonomske fakultete Univerze v Ljubljani, avtorica predloženega dela z naslovom Analiza stanja in poslovnih priložnosti uporabe masovnih podatkov v velikih slovenskih podjetjih, pripravljenega v sodelovanju s svetovalcem prof. dr. Alešem Groznikom,

IZJAVLJAM

1. da sem predloženo delo pripravila samostojno;
2. da je tiskana oblika predloženega dela istovetna njegovi elektronski obliki;
3. da je besedilo predloženega dela jezikovno korektno in tehnično pripravljeno v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani, kar pomeni, da sem poskrbela, da so dela in mnenja drugih avtorjev oziroma avtoric, ki jih uporabljam oziroma navajam v besedilu, citirana oziroma povzeta v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani;
4. da se zavedam, da je plagiatorstvo – predstavljanje tujih del (v pisni ali grafični obliki) kot mojih lastnih – kaznivo po Kazenskem zakoniku Republike Slovenije;
5. da se zavedam posledic, ki bi jih na osnovi predloženega dela dokazano plagiatorstvo lahko predstavljalo za moj status na Ekonomski fakulteti Univerze v Ljubljani v skladu z relevantnim pravilnikom;
6. da sem pridobila vsa potrebna dovoljenja za uporabo podatkov in avtorskih del v predloženem delu in jih v njem jasno označila;
7. da sem pri pripravi predloženega dela ravnala v skladu z etičnimi načeli in, kjer je to potrebno, za raziskavo pridobila soglasje etične komisije;
8. da soglašam, da se elektronska oblika predloženega dela uporabi za preverjanje podobnosti vsebine z drugimi deli s programsko opremo za preverjanje podobnosti vsebine, ki je povezana s študijskim informacijskim sistemom članice;
9. da na Univerzo v Ljubljani neodplačno, neizključno, prostorsko in časovno neomejeno prenašam pravico shranitve predloženega dela v elektronski obliki, pravico reproduciranja ter pravico dajanja predloženega dela na voljo javnosti na svetovnem spletu prek Repozitorija Univerze v Ljubljani;
10. da hkrati z objavo predloženega dela dovoljujem objavo svojih osebnih podatkov, ki so navedeni v njem in v tej izjavi.

V Ljubljani, dne _____

Podpis študentke: _____

KAZALO

UVOD	1
1 MASOVNI PODATKI.....	3
1.1 Razvoj masovnih podatkov	4
1.2 Metode obdelave masovnih podatkov	6
1.3 Tehnike analiziranja masovnih podatkov	7
1.4 Vir masovnih podatkov	8
1.5 Tveganja masovnih podatkov	9
2 POSLOVNE PRILOŽNOSTI V POVEZAVI Z MASOVNIMI PODATKI	10
2.1 Prilagojeno trženje.....	11
2.2 Prilagojene cene	12
2.3 Nižji stroški poslovanja in izvajanje napovednih finančnih modelov.....	12
2.4 Izboljšana podpora uporabnikom	12
2.5 Masovni podatki kot konkurenčna prednost	13
2.6 Uporaba masovnih podatkov v Sloveniji	13
2.7 Primeri uporabe masovnih podatkov v globalnih podjetjih	14
3 MASOVNI PODATKI KOT DEL PODJETJA	17
3.1 Vključenost masovnih podatkov v strategijo podjetja	17
3.2 Prihodnost masovnih podatkov v podjetjih.....	19
3.3 Masovni podatki in analitika masovnih podatkov kot del prihodnosti podjetij	20
4 KVANTITATIVNA RAZISKAVA UPORABE MASOVNIH PODATKOV	24
4.1 Metodologija	25
4.2 Namen in cilji raziskave	28
4.3 Raziskovalna vprašanja.....	29
4.4 Analiza	29
5 DISKUSIJA	45
5.1 V Sloveniji analizirajo masovne podatke velika podjetja, ki izhajajo iz storitvene in proizvodne dejavnosti.....	46
5.2 Masovne podatke pozna večina velikih podjetij v Sloveniji, veliko jih te podatke tudi obdeluje	47

5.3	Ne morem trditi, da se velika podjetja v Sloveniji ne poslužujejo tehnik tekstovnega rudarjenja in napovedne analitike.....	47
5.4	Uporaba masovnih podatkov predstavlja enega izmed poslovnih procesov, za katere skrbijo v velikih podjetjih v Sloveniji	48
SKLEP		51
LITERATURA IN VIRI		53
PRILOGE		57

KAZALO TABEL

Tabela 1:	Storitve v obdobju spleta 1.0 in nadomestne storitve v obdobju spleta 2.0	5
Tabela 2:	Primerjava različnih tehnik analize masovnih podatkov.....	8
Tabela 3:	Najpogostejši viri masovnih podatkov	9
Tabela 4:	Primer podjetij, ki v poslovanju uporabljajo masovne podatke	15
Tabela 5:	Seznam razvijajočih se raziskovalnih tehnik in njihova področja zanimanja ...	23
Tabela 6:	Prikaz odgovorov respondentov glede panoge, v katero z vidika dejavnosti sodi podjetje, v katerem so zaposleni	31
Tabela 7:	Prikaz odgovorov respondentov glede morebitne uporabe masovnih podatkov v prihodnosti v njihovem podjetju.....	33
Tabela 8:	Obdelava masovnih podatkov v podjetju respondenta.....	34
Tabela 9:	Odgovori respondentov glede vključenosti masovnih podatkov v strategijo njihovega podjetja	34
Tabela 10:	Odgovori respondentov glede merjenja učinka uporabe masovnih podatkov	35
Tabela 11:	Rezultati meritev učinka uporabe masovnih podatkov	35
Tabela 12:	Odgovori glede uporabe tekstovnega rudarjenja in napovedne analitike v podjetjih anketiranih.....	36
Tabela 13:	Prikaz odgovorov respondentov glede področja vlaganje finančnih sredstev v okviru masovnih podatkov	38
Tabela 14:	Prikaz odgovorov respondentov glede področij, na katera želijo vpliva v okviru masovnih podatkov	39
Tabela 15:	Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in spremenljivko »Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?«	42
Tabela 16:	Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in spremenljivko »Opažate pozitiven ali negativen učinek na poslovanje?«	43
Tabela 17:	Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in	

spremenljivko »Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?	43
Tabela 18: Pearsonov korelacijski koeficient med spremenljivko »Ali v vašem podjetju obdelujete t. i. masovne podatke?« in spremenljivko »... pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem ...?	44
Tabela 19: Prikaz odgovorov respondentov glede pomanjkanja znanja na področju masovnih podatkov	45
Tabela 20: Prikaz odgovorov respondentov glede izzivov, s katerimi se soočajo	45

KAZALO SLIK

Slika 1: Analiza masovnih podatkov v slovenskih podjetjih za leto 2015 (v %).....	14
Slika 2: Razvoj odnosa med podatki in strategijo	18
Slika 3: Pregled rasti količine podatkov skozi leta (v zettabajtih)	19
Slika 4: Tehnologije industrije 4.0	21
Slika 5: Spletno iskanje na globalni ravni, po ključnih besedah v angleškem jeziku od 2010 do 2018.....	22
Slika 6: Iskane sorodne ključne besede, povezane s ključno besedo »Big Data«.....	23
Slika 7: Odstotek respondentov glede na spol	30
Slika 8: Odstotek respondentov na določenem položaju v podjetju: uprava, vodja in izvedbena funkcija	30
Slika 9: Prikaz odgovorov respondentov glede poznavanje izraza »masovni podatki« (v %).....	32
Slika 10: Preizkušanje hipoteze enake verjetnosti	32
Slika 11: Preizkušanje hipoteze enake verjetnosti za vprašanje Q5a: Tekstovno rudarjenje	36
Slika 12: Preizkušanje hipoteze enake verjetnosti za vprašanje Q5b: Napovedna analitika.....	37
Slika 13: Preizkušanje hipoteze enake verjetnosti za Q10: V kateri del vašega podjetja na področju masovnih podatkov ste vložili največ?	39
Slika 14: Grafični prikaz odgovorov respondentov, na katera področja želijo vplivati z masovnimi podatki	41
Slika 15: Preizkušnja hipoteze enake verjetnosti za vprašanje Q11: Na katera področja vašega podjetja poskušate vplivati z masovnimi podatki?	41

KAZALO PRILOG

Priloga 1: Elektronsko sporočilo z vabilom k izpolnitvi spletne ankete	1
Priloga 2: Anketni vprašalnik: Analiza uporabe masovnih podatkov	2

SEZNAM KRATIC

ang. – angleško

BIRCH – algoritem združevanja podatkov

Big Data – masovni podatki

CLARA – ena izmed tehnik rudarjenja podatkov

CLARANS – ena izmed tehnik rudarjenja podatkov

EU – Evropska unija

EUR – evro

FLOT – orodje za prikaz podatkov

GGPLOT2 – eden od glavnih paketov R za prikaz in oblikovanje statističnih podatkov

GPS – (ang. Global Positioning System); globalni sistem pozicioniranja

HTML – (ang. Hyper Text Markup Language); glavni programski jezik za oblikovanje spletnih strani

IoT – (ang. Internet of things); internet stvari

IT – (ang. Information technology); informacijska tehnologija

LLE – (ang. Locally Linear Embedding); linearna tehnika analize lokalnih podatkov

LOGML – standardna oblika zapisa (besedilnih) dnevnikov, ki se zbirajo ob uporabniškem dostopu do spletnega strežnika

LTSA – nelinearna tehnika analize podatkov

PCA – (ang. Principal Component Analysis); analiza glavne komponente

RFID – (ang. Radio Frequency Identification Technology); tehnologija radio frekvenčne identifikacije

RS – Republika Slovenija

SMS – (ang. Short Message Service); kratko tekstovno sporočilo

SURS – Statični urad Republike Slovenije

SQL – (ang. Structured Query Language); strukturiran povpraševalni jezik za delo s podatkovnimi bazami

WSN – (ang. Wireless Sensor Networks); brezžično omrežje senzorjev

ZDA – Združene države Amerike

UVOD

Digitalno poslovanje je v zadnjih letih pridobilo pomembno vlogo v globalni ekonomiji. Govorimo lahko o digitalni ekonomiji, ki jo opredelimo kot vrsto družbenih, ekonomskih in kulturnih aktivnosti, ki se izvajajo in so povezane s spletom ter z uporabo informacijsko-komunikacijskih tehnologij (Zupan, 2016a, str. 27). Z uporabo digitalnih storitev in izdelkov, kot so pametni mobilni telefoni, spletne strani, pametne naprave, senzorji ipd., se iz dneva v dan ustvarja več podatkov. Z vse večjo količino podatkov pa se srečamo s tako imenovanimi masovnimi podatki (ang. Big Data), ki so že lep čas predmet obravnave različnih akademskih, strokovnih in raziskovalnih inštitucij. Masovni podatki so velika oziroma, rečemo lahko, ogromna in kompleksna zbirka podatkov, ki obstajajo že več desetletij. V začetku leta 2010 se pojavi termin »Big Data« oziroma »masovni podatki«, ki opisuje omenjene ogromne in kompleksne baze podatkov. Definicija masovnih podatkov sicer ne obstaja, obstajajo pa lastnosti, ki določajo termin. Te lastnosti so količina, raznolikost, več virov in nestanovitnost (Finlay, 2014, str. 13).

V raziskovalni organizaciji McKinsey so leta 2011 zapisali, da pričakujejo 40 % letno rast podatkov na globalni ravni, kjer pa je bilo že v tistem letu 30 milijard deljenih vsebin med uporabniki zgolj družabnega omrežja Facebook (Manyika in drugi, 2011, str. 18–19). Potreba po hranjenju, obdelavi in analizi podatkov postaja vedno večja. V Združenih državah Amerike (v nadaljevanju ZDA) podjetja skupno, ne glede na dejavnost, hranijo približno 100 terabajtov podatkov. Podatki, ki se hranijo, so lahko v različnih oblikah zapisa: slika, video, zvočni posnetek in tekstovni zapis. Skoraj v vseh dejavnostih predstavlja največjo količino podatkov tekstovni zapis. Sledijo video in zvočni posnetki v komunikacijskih, medijih, vladnih in izobraževalnih dejavnostih. Največ slik pa ustvarijo in hranijo na področju zdravstvene in gradbeniške dejavnosti (Manyika in drugi, 2011, str. 19–20).

Obdobje, v katerem živimo, je vse bolj preprejeno z napravami, ki naše življenje iz dneva v dan bolj digitalizirajo in s tem omogočajo sledenje našim digitalnim »sledem« oziroma aktivnostim. Te sledi v resnici predstavljajo podatke, ki jih ustvarjamo mi, uporabniki, ko pregledujemo spletne strani, ko uporabljamo pametne naprave ipd. Kot predstavljajo Akter, Wamba, Gunasekaran, Dubey in Childe (2016, str. 124), imajo analitične zmožnosti masovnih podatkov statistično značilni vpliv na uspešno in učinkovito poslovanje podjetja. Na drugi strani pa podatke ustvarjajo tudi podjetja sama, ko izvajajo poslovne funkcije na digitaliziran način. Iz upravljanja masovnih podatkov izhajajo številne poslovne priložnosti za podjetja. Podjetja, ki v svojem poslovanju izvajajo analize masovnih podatkov, lahko izvajajo prilagojeno trženje ne zgolj z vidika prilagajanja vsebine končnemu uporabniku, temveč tudi z vidika prilagajanja cen izdelkov. Podjetja lahko na osnovi analize masovnih podatkov tudi znižajo stroške lastnega poslovanja oziroma optimizirajo poslovne procese, da ti postanejo hitrejši, kakovostnejši oziroma učinkovitejši. Priložnosti pa se ne kažejo zgolj z vidika analize »notranjih« masovnih podatkov, ki so pretežno podatki, ki jih podjetja ustvarjajo sama z lastnimi procesi in sistemi. Druge priložnosti so priložnosti, ki se nanašajo na »zunanje« poslovanje podjetja oziroma ustvarjanje konkurenčne prednosti na podlagi

zbranih in analiziranih podatkov, personaliziranega oglaševanja ipd. Poslovne prednosti masovnih podatkov, kot navajajo Alharthi, Krotov in Bowman (2017, str. 286), se kažejo predvsem v izboljšani uporabniški izkušnji, optimiziranem internem poslovanju podjetja ter ključno v izboljšanju dobička in konkurenčnosti na trgu. Iz statističnega pregleda 2016 je razvidno, da je v letu 2015 masovne podatke v Sloveniji že analiziralo 11 % podjetij iz različnih gospodarskih dejavnosti (z vsaj 10 zaposlenimi) (Zupan, 2016b).

Iz raziskave, ki so jo izvedli Henke in drugi (2016), izhaja, da je v obdobju od leta 2011 do leta 2016 največji napredek na področju koriščenja masovnih podatkov na področju, kjer so aktivna podjetja, ki nudijo storitveno dejavnost in uporabljajo lokacijske podatke, ter tudi v maloprodaji. V nedavni raziskavi je bilo tudi ugotovljeno, da je možno že v okvirno enem desetletju ne le povrniti, temveč podvojiti prihodke in prihranke v poslovanju podjetja od naložb v podatkovne in analitične zmogljivosti (Henke in drugi, 2016, str. 29–30).

Podjetja, ki danes veljajo za ena od najbogatejših ter naprednejših, temeljijo na poslovnih modelih, kjer je eden od ključnih dejavnikov prav podatek. Že več let se govori o visoki vrednosti podatka. Menim, da se šele zadnjih nekaj let zares zavedamo, kako korenito lahko ti vplivajo na poslovne rezultate podjetja. Kot navajajo Fan, Han in Liu (2014, str. 296), se vedno več organizacij, predvsem korporacij, odloča za vodenje na podlagi podatkovne informiranosti z namenom ustvarjanja in nudenja bolj usmerjenih storitev, zniževanja tveganj in izboljšanja učinkovitosti poslovanja podjetja.

V prihodnosti pa je pričakovati še večje, kompleksnejše podatke. Kot je zapisal de Waal-Montgomery (2015), je pričakovana letna rast podatkov 40 %, kot sem tudi že omenila na začetku. Takšno rast lahko pričakujemo predvsem zaradi interneta stvari, kjer gre za globalno mrežo povezanih naprav, ki so zmožne medsebojno komunicirati (Lee & Lee, 2015, str. 431). Hitri dostop do podatkov in informacij, ki jih ustvarjajo naprave interneta stvari, omogoča ustvarjanje visoko inovativnih in učinkovitih storitev, te pa omogočajo definiranje različnih poslovnih modelov podjetij (Metallo, Agrifoglio, Schiavone & Mueller, 2018).

Zanima me, ali se tudi v Sloveniji velika podjetja zavedajo priložnosti, kaj že počnejo na področju masovnih podatkov, kaj načrtujejo ter s kakšnimi izzivi se srečujejo pri upravljanju podatkov. Tako sem se odločila za analizo tematike v okviru magistrskega dela in določila namen naloge kot interes za proučitev ozaveščenosti, obstoječega stanja in naklonjenosti uporabi masovnih podatkov v velikih slovenskih podjetjih. Hkrati je namen naloge tudi prikazati nabor poslovnih priložnosti, ki izhajajo iz uporabe masovnih podatkov in bi lahko vzbudile zanimanje tudi velikih podjetij v Sloveniji.

Cilj magistrskega dela je preveriti, ali obstajajo v velikih slovenskih podjetjih na področju masovnih podatkov neizkoriščene poslovne priložnosti, ki se lahko kažejo v različnih oblikah.

V prvem delu naloge je najprej na kratko predstavljen pojem, skupaj z zgodovino, metodami obdelave, tehnikami ter tveganji in izzivi uporabe masovnih podatkov. Po začetni podrobnejši seznanitvi bralca s proučevanim področjem masovnih podatkov naloga preide na poslovne priložnosti, ki izhajajo iz uporabe masovnih podatkov za podjetja. Poleg teoretičnih izhodišč in domnev so zajeti tudi praktični, realni primeri iz znanih podjetij, kot so Walt Disney, Etihad Airways ipd., ki masovne podatke uporabljajo na dnevni ravni poslovanja in z njihovo uporabo zasledujejo cilj optimizacije poslovanja, tako z vidika nižanja stroškov poslovanja kot tudi ustvarjanja zadovoljnih strank, ki posledično pomenijo tudi več prihodkov in dobička za podjetje. Zgolj zavedanje, da podjetje s poslovanjem ustvarja podatke, pa še ne zadošča, da bi podjetja te podatke uporabljala tudi v poslovanju. Tako v zadnjem delu teoretičnega sklopa naloge sledi še pregled pomembnosti vključenosti masovnih podatkov kot področja ene izmed poslovnih funkcij podjetja v strategijo podjetja.

Zadnji del magistrskega dela je empiričen. Sledi poglavje kvantitativne raziskave, kjer najprej povzamem zastavljeni cilj, namen, raziskovalno vprašanje in eksperimentalni hipotezi, nato sledi opis uporabljene metodologije ter na koncu še analiza pridobljenih podatkov ter diskusija. V analitičnem delu se osredotočam na rezultate spletne ankete, izvedene v velikih podjetjih v Sloveniji. Sledi diskusija, kjer ugotovitve iz empiričnega dela naloge primerjam s teoretičnimi izhodišči povzetih znanstvenih in strokovnih avtorjev, ki so bila tudi osnova za oblikovanje vprašanj v spletni anketi. V sklepnem delu pa je predstavljen skupek ugotovitev, spoznanj ter priporočil za nadaljnje raziskovanje področja.

1 MASOVNI PODATKI

Velike in kompleksne podatkovne zbirke obstajajo že več desetletij, tako pri masovnih podatkih v bistvu ne gre za popolnoma novo ali neznano področje. Za razliko od podatkovnih zbirk masovni podatki predstavljajo ogromne količine obsežnih in kompleksnih podatkov (Finlay, 2014, str. 13). Kot navaja Finlay (2014, str. 13), definicija masovnih podatkov ni usklajena, navaja pa nekaj ključnih lastnosti, ki opredeljujejo masovne podatke, kot so: količina, raznolikost, hitrost in več virov:

- Z vidika količine: podatkovne zbirke, ki so tako velike, da jih je težko upravljati na povprečnem osebem računalniku, razumemo kot masovne podatke. Če izhajamo iz količine zajetih podatkov, potem lahko o masovnih podatkih govorimo, ko gre za količino enega ali več terabajtov podatkov. Pri določenih virih podatkov pa govorimo tudi o količini, izraženi v petabajtih.
- Z vidika raznolikosti: masovni podatki so sestavljeni iz različnih tipov in virov podatkov, ki so lahko strukturirani ali nestrukturirani. Strukturirane podatke je lažje prikazati in organizirati v primerjavi z nestrukturiranimi, kjer je organiziranost podatkov zahtevnejša. Med strukturirane podatke sodijo: prihodki, starost, spol itd., med nestrukturirane pa sodijo: elektronska sporočila, spletni dnevniki, spletne strani, prepisi in telefonski pogovori.

- Z vidika hitrosti: poznamo različne tipe podatkov glede na njihovo hitrost spreminjanja. Poznamo statične oziroma matične podatke, podatke, ki se spreminjajo občasno in počasi, ter dinamične, nenehno spreminjajoče se podatke.

Matični podatki se načeloma ne spreminjajo oziroma zelo redko. Med matične podatke lahko štejemo: spol, državljanstvo, kraj rojstva itd. Med podatke, ki se spreminjajo občasno in tudi počasi, sodijo: naslov prebivališča, ime delodajalca in število otrok. Nenehno spreminjajoči se podatki so najzahtevnejši podatki z vidika zajema, obdelave in analize. Podatki, ki jih lahko uvrščamo v to skupino, so: vrsta glasbe, ki jo nekdo posluša, in hitrost vožnje avtomobila. Slednji tip podatkov sodi v skupino masovnih podatkov, za katero je značilna kompleksnost, kot sem že zapisala na začetku (Finlay, 2014, str. 13).

V primeru več virov: določeni podatki se ustvarjajo zgolj znotraj organizacije, drugi so ustvarjeni tudi zunaj nje. Masovni podatki navadno zajemajo tako zunanje kot tudi notranje podatke. Med zunanje podatke podjetja sodijo: kreditna poročila, informacije o popisih, GPS-podatki, spletne strani itd. Dodatno je pomembno opozoriti, da podjetja na slednje podatke navadno nimajo vpliva, kar lahko povzroča tudi nekatere izzive na področju kakovosti podatkov (Finlay, 2014, str. 13–14).

1.1 Razvoj masovnih podatkov

Zbiranje podatkov ima precej daljšo zgodovino, kot si verjetno predstavljamo. Zbiranje in hranjenje podatkov obstaja že od zgodnjih 50. let 20. stoletja, medtem ko se je pojem masovni podatki pojavil šele pred kratkim (Lee, 2017, str. 296). Kot sem omenila že v začetnem delu naloge, se je pojem pričel pogosteje uporabljati približno pred osmimi leti, ko je tudi pričela nastajati vse večja potreba po hranjenju in obdelavi podatkov, kot prikazujemo v nadaljevanju, pa segajo začetki masovnih podatkov še nekaj več let nazaj, v obdobje konca 90. let 20. stoletja.

Kot meni Lee (2017, str. 296), je bila rast ustvarjenih podatkov v obdobju med začetkom 50. in sredino 90. let 20. stoletja precej počasna. Počasna rast količine podatkov je bila predvsem posledica visokih stroškov računalniške opreme, hranjenja podatkov in podatkovnih mrež v tistem času. Razvoj svetovnega spleta konec 90. let je botroval nenadnemu poskoku v nastajanju podatkov in razvoju masovnih podatkov ter analitiki masovnih podatkov (Lee, 2017, str. 296). Masovni podatki so v razvoju od leta 1994 do 2015 prešli tri večja obdobja: masovni podatki 1.0 (1994–2004), masovni podatki 2.0 (2005–2014) in masovni podatki 3.0 (2015–) (Lee, 2017, str. 296).

- Masovni podatki 1.0

V obdobju med 1994 in 2004 se je pričela pojavljati spletna prodaja, ki je prispevala k ustvarjanju vedno večje količine podatkov. V tem obdobju so podatke ustvarjala pretežno podjetja, uporabniške vsebine in podatki v tem obdobju še niso bili tako pogosti. Lee (2017,

str. 296) navaja tudi, da so se v tem obdobju pojavile tehnike spletnega rudarjenja z namenom analiziranja uporabniškega vedenja.

– Masovni podatki 2.0

Lee (2017, str. 296) navaja, da se je obdobje masovnih podatkov 2.0 pojavilo vzporedno z razvojem spleta 2.0, zato masovne podatke 2.0 umešča v obdobje med letoma 2005 in 2014.

Z izrazom splet 2.0 (ang. Web 2.0) je O'Reilly poskušal definirati prelomnico v načinu uporabe spletnih storitev in načinu komunikacije na spletu, ki se je spreminjala v tistem obdobju zaradi napredka v tehnologiji. Pojavilo se je veliko kritikov, ki so bili mnenja, da je izraz splet 2.0 le marketinška fraza in v resnici nima večjega pomena, kot navaja O'Reilly (2007). Tako je poskušal popisati obstoječe spletne storitve in funkcionalnosti na način, da bi jih lahko opredelil kot storitve spleta 1.0 in spleta 2.0 ter tako prikazal razlike, ki so se pojavile v tem obdobju, kot je razvidno iz tabele 1.

Tabela 1: Storitve v obdobju spleta 1.0 in nadomestne storitve v obdobju spleta 2.0

Splet 1.0	Splet 2.0
DoubleClick	Google AdSense
Ofoto	Flickr
Akamai	BitTorrent
mp3.com	Napster
Britannica Online	Wikipedia
lastna spletna stran (ang. personal websites)	pisanje spletnega dnevnika (ang. blogging)
evite	upcoming.org and EVDB
poimenovanje domen (ang. domain name speculation)	optimizacija spletne strani za spletne iskalnike (ang. search engine optimization)
ogledi strani (ang. page views)	cena na klik (ang. cost per click)
»strganje« zaslona (ang. screen scraping)	spletne storitve (ang. web services)
objave (ang. publishing)	sodelovanje (ang. participation)
sistem za upravljanje vsebine (ang. content management system)	wikis
directories (taxonomy)	označevanje (ang. tagging („folksonomy“))
lepljivost (ang. stickiness)	sindikacija (ang. syndication)

Vir: O'Reilly (2007, str. 18).

Kot navaja O'Reilly (2007), je v obdobju spleta 2.0 vsaka zahtevnejša in naprednejša spletna storitev slonela na podatkovnih bazah in programski jezik SQL naj bi v tem času postajal nov HTML. Upravljanje podatkov v tem obdobju za podjetja postane ena izmed ključnih aktivnosti (O'Reilly, 2007, str. 27).

V obdobju spleta 2.0 so se pojavila družabna omrežja, ki so povsem spremenila način komunikacije in dela tako končnih uporabnikov kot podjetij (Lee, 2017, str. 296). Družabna omrežja so tako kot tudi danes vzpodbujala komunikacijo uporabnikov prek spleta. Z uporabo storitev za komuniciranje znotraj družabnih omrežij pa uporabniki ustvarjajo tudi

podatkovne zapise, kar je privedlo do spletne analitike in analitike družabnih omrežij. Razvile so se različne tehnike analiziranja podatkov. Ena izmed tehnik je rudarjenje podatkov, kjer lahko rudarimo po podatkih na različne načine in po različnih vsebinah (Lee, 2017, str. 296–297). Poleg tehnike rudarjenja sta se pojavili tudi tehnika analize čustvenega odziva (ang. Sentiment analysis) ter analitika družbenih omrežij, ki je predvsem proces merjenja in spremljanja strukture, povezav in drugih lastnosti družabnih omrežij (Lee, 2017, str. 297).

– Masovni podatki 3.0

Masovni podatki 3.0 zajemajo vse podatke, ki so zajeti med masovnimi podatki 1.0 in 2.0, ter dodatno zajemajo tudi podatke v obliki slik, besedil, zvoka in videa, ki so ustvarjeni ob uporabi izdelkov tako imenovanega interneta stvari (ang. Internet of Things). O masovnih podatkih 3.0 govorimo od leta 2015 naprej, ko so se tudi pričeli masovno pojavljati izdelki interneta stvari. Izdelki in senzorji, ki sodijo med izdelke interneta stvari, imajo unikatne identifikatorje, ki omogočajo deljenje podatkov in soustvarjanje več deležnikov prek interneta brez kakršnegakoli človeškega posega. Zajem tovrstnih podatkov predstavlja dodatne, nove tipe podatkov in vodi do novih oblik obdelave in analize podatkov, do analitike pretakanja podatkov. Analitika pretakanja podatkov ima velik potencial v številnih industrijah, kjer se pojavlja prenos podatkov, ki ga ustvarjajo ljudje, naprave ali senzorji (ang. Streaming analytics) (Lee, 2017, str. 297–298). Več o prihodnosti in priložnostih, predvsem z vidika poslovnih priložnosti, je predstavljeno v nadaljevanju naloge.

1.2 Metode obdelave masovnih podatkov

Z različnimi metodami obdelave masovnih podatkov lahko iz velike količine podatkov pridobimo informacije z visoko vrednostjo. Metode za obdelavo masovnih podatkov so (Yaqoob in drugi, 2016, str. 1238):

– Hash algoritem (ang. Hashing)

S pomočjo Hash algoritma lahko z učinkovito metodo pridobimo podatke z diska brez uporabe indeksne strukture. Gre za metodo preslikave (ang. mapping), ki jemlje vrednost kot vnos in njeno vrednost pretvori v ključ. Vrednost ključa določa lokacijo podatkov. Hash datoteke tako hranijo podatke v obliki seznama. Poznamo statično in dinamično izvajanje Hash algoritma. Pri izvajanju statičnega načina funkcija vedno izračuna enako lokacijo, ko je tudi zagotovljena vrednost ključa. Slabost statične funkcije je, da kadar se naenkrat pojavi velika količina podatkov, se sezname ne zmanjšujejo dinamično. Pri dinamični funkciji Hash algoritma se sezname dodajajo dinamično in so hkrati odstranjeni glede na potrebe. Uporaba metode je najbolj priporočljiva, kadar so podatki diskretni in naključni.

- Bloomov filter (ang. Bloom Filter)

Pri metodi uporabe Bloomovega filtra gre za hranjenje vrednosti v obliki Hash zapisa, namesto hranjenja podatkov z uporabo bitne matrike. Prednost predstavljene metode je predvsem v hitrosti izvajanja poizvedb in nizki porabi prostora.

- indeksacija (ang. Indexing)

Metoda indeksacije se uporablja za hitro lociranje podatkov med velikimi in kompleksnimi količinami podatkov. Izvajanje ročnih pregledov ni smotno, zgolj indeksacija z velikim pretokom lahko zagotovi rezultate hranjenja masovnih podatkov. Poznamo različne tipe indeksacij: semantično indeksacijo, indeksacijo datotek, r-drevesno indeksacijo (ang. r-tree indexing) in druge. Edina slabost izvajanja metod indeksacije so visoki stroški priklica, tako je tudi trenutno zelo aktualno področje razvoja učinkovite in relativno ugodne metode indeksacije.

- Vzporedno računalništvo (ang. Parallel computing)

Metoda vzporednega računalništva omogoča koriščenje več virov (računalniki, programska oprema, ipd.) hkrati z namenom zaključitve določene naloge. V primeru kompleksnejših nalog ali problemov so le ti razdeljeni na več izvedbenih delov med več računalnikov, sistemov ali programskih rešitev, kjer se vsak izmed porazdeljenih delov izvaja hkrati.

1.3 Tehnike analiziranja masovnih podatkov

Kot navajajo Yaqoob in drugi (2016, str. 1240), so za izvedbo ustrezne analize velike količine podatkov v relativno kratkem času potrebne posebne tehnike obdelave masovnih podatkov. Zaenkrat je na voljo omejeno število tehnik, ki omogočajo analizo masovnih podatkov. Te tehnike so:

- podatkovno rudarjenje (ang. Data Mining),
- analiza družabnih omrežij (ang. Social Network Analysis),
- spletno rudarjenje (ang. Web Mining),
- strojno učenje (ang. Machine Learning),
- pristopi vizualizacije (ang. Visualization Approaches) in
- optimizacije (ang. Optimization Methods).

Različne tehnike uporabljamo predvsem z vidika namena uporabe podatkov ter pričakovanj, kaj želimo na podlagi podatkov in z izbiro določene tehnike analize ugotoviti. Z namenom lažjega razlikovanja med posamezno tehniko v nadaljevanju sledi primerjalna tabela (tabela 2), ki nazorno prikazuje vsako izmed tehnik z vidika namena uporabe, področja uporabe, uporabljenih algoritmov ter orodij, ki so na voljo za izvedbo določene analize.

Tabela 2: Primerjava različnih tehnik analize masovnih podatkov

Tehnike analiz	Namen	Uporaba	Algoritmi	Razpoložljiva orodja
Podatkovno rudarjenje (ang. Data Mining)	Najti skladne vzorce ali sistematične vezi med spremenljivkami	Biomedicina, zdravstvo	Metoda za razvrščanje v skupine (ang. K-mean), sintetična in analitična metoda za reševanje logičnih nalog (ang. Fuzzy C-Mean), CLARA, CLARANS, BIRCH	Excel, Rapid – I, Rapidminer – R, KNMINE, Weka/Pentaho
Analiza družabnih omrežij (ang. Social Network Analysis)	Najti družbene vezi in pogoje za oblikovanje teorije mreženja	Antropologija, družabna omrežja	PCA, LTSA, LLE, Autoencoder	Cytoscape, Gephi, Cuttlefish, MeerKat
Spletno rudarjenje (ang. Web Mining)	Za odkrivanje vzorcev v načinu uporabe velikih spletnih repozitorijev	E-učenje, digitalne knjižnice, e-uprava	LOGML, Apriori	KXEN, LIONsolver, Dataiku
Strojno učenje (ang. Machine Learning)	Za razvoj »samostojnega« računalniškega vedenja na podlagi empiričnih podatkov	Zdravstvo, pomoč uporabnikom	Prepoznavanje vezi (ang. Pattern recognition), umetna nevronska mreža (ang. Artificial neural Networks)	Weka, Scikit-Learn, PyMc, Shogun
Pristopi vizualizacije (ang. Visualization Approaches)	Predstavitev podatkov in znanja z uporabo grafov	Bančništvo, proizvodnja	FLOT, GGPlot2	Data wrapper, Highcharts JS, MAPBox
Metode optimizacije (ang. Optimization Methods)	Za reševanje kvantitativnih težav	Znanost družabnih omrežij, računalniška biologija	Zmanjšanje, paralelizacija (ang. Parallelization), simulirano žarjenje (ang. Simulated annealing), kvantno žarjenje (ang. Quantum annealing), optimizacija rojstva (ang. Swarm optimization)	Metlab

Vir: Yaqoob in drugi (2016, str. 1239).

1.4 Vir masovnih podatkov

Podjetja, predvsem vodstveni delavci, se srečujejo z velikimi spremembami na področju vodenja podjetja. Tem spremembam botruje nastanek »revolucije« na področju vedno večjega števila podatkov. Z vse hitrejšim razvojem tehnologij nastaja vedno več podatkov,

ki omogočajo vodenje na vedno bolj informiran način. Na podlagi vedno večjega števila podatkov, ki nastajajo na poslovnih sistemih oziroma virih masovnih podatkov, tako podjetja pridobivajo vedno več možnosti zbiranja, obdelave in analize vse večjega števila podatkov.

Podjetja zbirajo podatke iz različnih lastnih in tudi tujih virov, na podlagi katerih lahko izvajajo analize podatkov. Raguseo (2018, str. 189-190) je na podlagi raziskave zapisal vire, ki jih podjetja uporabljajo in so predstavljeni tudi v tabeli 3. Ugotovil je, da podjetja, ki so bila vključena v vzorec, največ uporabljajo spletne vire in spletne vsebine kot vir masovnih podatkov, ki pa mu sledijo prodajna mesta in podatki pametnih merilnikov.

Tabela 3: Najpogostejši viri masovnih podatkov

Vir masovnih podatkov v podjetjih:
Spletni vir/spletne vsebine
Podatki iz prodajnih mest ali drugi transakcijski podatki
Podatki pametnih merilnikov
Podatki video vsebin
Podatki slikovnih materialov
Drugi viri tekstovnih virov
Objave družbenih omrežij
RFID sistemski podatki
Podatki zvočnih posnetkov
Podatki klikov
GPS-podatki
Objave spletnih dnevnikov/blogov
Objave mikroblogov (Twitter)

Vir: Raguseo (2018, str. 191).

1.5 Tveganja masovnih podatkov

Z vidika podjetja in uporabe masovnih podatkov v poslovne namene se tveganja kažejo na različnih področjih. Kot navajajo Sivarajah, Kamal, Irani in Weerakkody (2017, str. 265), se tveganja kažejo kot karakteristike masovnih podatkov. Določena tveganja se kažejo pri obstoječih analitičnih metodah in modelih obdelave, določena pa z omejitvijo obstoječih sistemov obdelave podatkov v podjetju (Mantelero, 2017, str. 584–602). Tveganja so prisotna na področju analize masovnih podatkov, kjer je prvo takšno tveganje, izraženo v napaki, pristranskost pri iskanju vzorcev (Pohorec, 2015a, str. 40). Napaka pristranskosti se lahko pojavi pri izvedbi analize. Pogosto v želji, da bi dosegli rezultate, tako vidimo le vzorce v podatkih, zanemarimo pa vzorce, ki zavračajo našo predpostavko raziskave (Pohorec, 2015a, str. 40). Obstajata še dve napaki, in sicer osredotočanje na nepomembne podatke, kjer se ukvarjamo s podatki, ki ne vodijo do zelenega cilja, oziroma se osredotočamo na napačne podatke, ter napaka napačne interpretacije vzročnosti in povezanosti (Pohorec, 2015a, str. 40).

Masovni podatki v praksi predstavljajo »goro« podatkov, kjer se lahko srečamo z izzivom, kako se lotiti analize takšne količine, da ne zaidemo v težave, kot so pristranskost, usmerjenost na napačen vzorec ipd. Kot navaja Pohorec (2015b, str. 41), ima vsaka analiza svoj cilj, zato iščemo najboljše odgovore na vprašanja, ki ta cilj opredeljujejo. Moderni pristopi analiziranja velikih količin podatkov temeljijo na kompleksni statistiki, matematiki ter mnogih drugih ozko usmerjenih metodah in pristopih (Pohorec, 2015b, str. 41). Izvajanja analiz masovnih podatkov v mnogih primerih temeljijo na podatkovnih jezerih (ang. Data Lake), kar predstavlja možnost vzpostavitve dialoga med različnimi viri podatkov in sestave celovitejše slike za določeno proučevano področje, hkrati pa tudi omogoča izvedbe ocene tveganj (Pohorec, 2015b, str. 42). Pohorec (2015b, str. 42) navaja, da je smotrno vzpostaviti t. i. peskovnik za masovne podatke, ki omogoča vzpostavitev podobnega okolja realnemu oziroma produkcijskemu z namenom izvajanja testiranj in preverjanj analiz pred dejansko obsežno analizo, ki zahteva večja finančna in časovna vlaganja.

Tveganje pri uporabi masovnih podatkov je lahko tudi za končne uporabnike spletnih storitev. S ciljem zaščite končnih uporabnikov v Evropski uniji (v nadaljevanju EU) so države članice Sveta Evrope podpisale Konvencijo o varstvu posameznika glede na avtomatsko obdelavo osebnih podatkov, ki je bila ratificirana leta 1994 (Svet Evrope, 1994). Kasneje je nastala tudi Splošna uredba o varstvu osebnih podatkov št. 2016/679 (Uredba (EU) 2016/679 Evropskega parlamenta in Sveta, 2016), ki je stopila v veljavo 25. maja 2018 (Uradni list Evropske unije, 2016, L119/65).

Dodatna tveganja in izzivi so prisotni tudi pri razumevanju in uporabi masovnih podatkov, pri načinu zajema najpomembnejših podatkov ob njihovem nastanku ter kako lahko dobljene podatke in informacije dostavimo dejansko najpomembnejšim subjektom v ustreznem času. Izziv pa je tudi, kako ustrezno hraniti vse zbrane podatke (Misra, Sharma, Gulia & Bana, 2014, str. 41).

Ne glede na vse predstavljene izzive in tveganja pa masovni podatki predstavljajo predvsem tudi veliko priložnosti v poslovanju, ki jih predstavljam v nadaljevanju naloge.

2 POSLOVNE PRILOŽNOSTI V POVEZAVI Z MASOVNIMI PODATKI

Uporaba podatkov v podjetjih z namenom zagotavljanja boljšega, bolj informiranega sprejemanja odločitev ter izvajanja poslovnih aktivnosti v zadnjih letih dramatično narašča (Miloslavskaya & Tolstoy, 2016, str. 300). Masovni podatki in analitika masovnih podatkov imata danes velik pomen v poslovanju, saj imata zelo močan vpliv na učinkovitost poslovanja podjetij in njihovo uspešnost (Wamba in drugi, 2017, str. 357). Zbiranje in analiza masovnih podatkov, ki nastajajo v vsakodnevnem okolju z uporabo in delovanjem različnih naprav, nudita možnosti za razvoj novih poslovnih priložnosti. Nove poslovne priložnosti za razvoj novih prilagojenih izdelkov in storitev, izboljšanje poslovanja z vidika

zniževanja stroškov in večjega učinka ter zagotavljanja večjega zadovoljstva strank. Ekonomske priložnosti, ki izhajajo iz uporabe masovnih podatkov, so ključne za poslovanje podjetij (Günther, Rezazade Mehrizi, Huysman & Feldberg, 2017, str. 191). Kot ugotavljajo Akter, Wamba, Gunasekaran, Dubey in Childe, imajo analitične zmožnosti masovnih podatkov (ang. Big Data Analytical Capabilities) statistično značilni vpliv na uspešno in učinkovito poslovanje podjetja (Akter, Wamba, Gunasekaran, Dubey & Childe, 2016, str. 124). Kot so zapisali na inštitutu McKinsey, obstaja pet načinov, kako masovni podatki ustvarjajo vrednost za podjetje (Manyika in drugi, 2011, str. 97–100):

- ustvarjajo lahko transparentnost zaradi širokega dostopa in uporabe,
- podjetjem omogočajo izvajanje raziskav in preizkusov na različnih področjih, kot sta na primer preizkušanje procesa sprememb in ugotavljanje možnosti za izboljšanje internih oziroma notranjih procesov poslovanja,
- uporaba podatkov za razvrščanje in poznavanje strank podjetja,
- podpora poslovnemu odločanju, ki omogoča vodenje na podlagi podatkov,
- na podlagi podatkov je možno prepoznati potrebe in zmožnosti za ustvarjanje novih izdelkov ali novih poslovnih modelov podjetja.

V nadaljevanju sledita podrobnejša predstavitev in pregled dejanskih možnosti uporabe masovnih podatkov, tudi na primerih iz prakse, in z njimi povezanih tehnik obdelave in analize masovnih podatkov.

2.1 Prilagojeno trženje

Z zbiranjem in analizo masovnih podatkov, ki izhajajo iz različnih virov, lahko podjetja prilagodijo ponudbo na način, da je ta prilagojena posameznemu obstoječemu ali potencialnemu uporabniku oziroma stranki ter njenim željam. Podjetje lahko prilagojeno ponudbo ponuja na več različnih načinov: prilagojeni kuponi za stranko, priporočila za določen izdelek ali posebna promocijska ponudba. Kot navaja Lee (2017, str. 300), naprednejša podjetja na področju financ že uporabljajo podatke z družabnih omrežij z namenom prepoznavanja tveganj pri odobritvi finančnih kreditov ter prepoznavanja finančnih potreb potencialnih strank. Na osnovi spoznanj iz masovnih podatkov, pridobljenih z družabnih omrežij, tako prilagodijo ponudbo potrebam strank.

Stranka lahko stopi v stik s podjetjem prek različnih komunikacijskih kanalov, ki jih nudi določeno podjetje. Kanali za stik so lahko prodajna mesta, klicni centri, spletni obrazci, elektronski naslovi itd. Podjetja tako zbirajo masovne podatke na ločenih kanalih, ki jih s pomočjo različnih metod lahko združijo, obdelajo in analizirajo, tako da imajo enoten pogled o aktivnostih strank povezanih z njihovim podjetjem, kjerkoli in kadarkoli. Tako je tudi U.S. Bank, ki je glavna banka v ZDA, razvila podatkovno analitiko, ki omogoča združevanje spletnih in fizičnih podatkov, na podlagi katerih izvajajo upravljanje odnosov s strankami (ang. Customer relationship management), kar jim je tudi izboljšalo določene rezultate poslovanja za 100 % (Lee, 2017, str. 300).

2.2 Prilagojene cene

Ne zgolj vsebinska prilagoditev izdelka ali storitve, na podlagi masovnih podatkov je možno tudi poznati pripravljenost strank za plačilo in tako prilagoditi ceno ponudbe. Kot navaja Lee (2017, str. 300), ameriško maloprodajno podjetje Sears uporablja masovne podatke z namenom prilagajanja cen in za program zvestobe strank. Na osnovi analize tako stranke prejmejo prilagojen vrednosti kupon, ki ga lahko unovčijo ob naslednjem nakupu, itd. (Lee, 2017, str. 300). Poleg možnosti prilagajanja cen in ponudbe je na osnovi masovnih podatkov možno tudi spremljati in izboljševati zadovoljstvo strank ter tako zvišati povpraševanje po izdelkih in storitvah v ponudbi, kar podrobneje predstavljam v nadaljevanju naloge.

2.3 Nižji stroški poslovanja in izvajanje napovednih finančnih modelov

Zmožnosti analitičnih metod, ki se uporabljajo pri analizi masovnih podatkov, imajo ključno vlogo, ko govorimo o rasti uspeha in učinkovitosti podjetja (Akter, Wamba, Gunasekaran, Dubey & Childe, 2016, str. 119). Podjetja z uporabo masovnih podatkov med drugim znižujejo operativne stroške poslovanja. Zniževanje stroškov je možno v različnih podjetjih oziroma podjetjih, ki se ukvarjajo z različnimi dejavnostmi. Kot primer Akter, Wamba, Gunasekaran, Dubey in Childe (2016, str. 119) na osnovi ugotovitev Srinivasa in Arunasalama navajajo primer, kako podjetja, ki so dejavna na področju medicine, znižujejo stroške poslovanja z zmanjšanjem odpadkov ter izboljšanjem kakovosti oskrbe pacientov, kar počnejo na osnovi ugotovitev, ki izhajajo iz analize masovnih podatkov njihovega poslovanja. Primer, ki ga navaja Lee (2017, str. 300) v dejavnosti maloprodaje, je podjetje Tesco, ki analizira podatke hladilnikov s ciljem znižanje porabe energije in posledično stroškov. Na osnovi analize podatkov so zgotovili, da letno prihranijo 25 milijonov dolarjev (Lee, 2017, str. 300).

Kot navajajo Gepp, Linnenluecke, O'Neill in Smith (2018, str. 104) je z vidika sprejemanja poslovnih odločitev na osnovi masovnih podatkov možno obravnavati več različnih področij. Zanimiva omenjena področja so na primer izboljšanje uporabniške izkušnje, ki jo obravnavam tudi v nadaljevanju magistrskega dela, ter tudi finančno področje, kjer je možno izdelati različne finančne modele, kot so: model finančnih stisk (ang. Finance distress model), model finančnih goljufij (ang. financial fraud model), napoved rasti vrednosti delnic ipd.. Gepp, Linnenluecke, O'Neill in Smith (2018, str. 104–105) prav tako navajajo, da se v največ študijah poslužujejo tehnik podatkovnega rudarjenja z namenom napovedovanja finančnih stisk, ki jih izdelujejo največkrat za obdobje enega leta, pri čemer pa, kot navajajo, bi lahko z uporabo omenjene tehnike izdelovali tudi napovedi za daljša časovna obdobja.

2.4 Izboljšana podpora uporabnikom

Podpora uporabnikom zahteva razumevanje in poznavanje tako uporabnikov kot tudi uporabniških potreb in morebitnih težav. Masovni podatki združujejo podatke iz različnih

komunikacijskih kanalov, kot so telefon, sporočila SMS, elektronska sporočila itd., ter tako omogočajo agentom, ki skrbijo za pomoč uporabnikom, popoln vpogled v uporabnikove pretekle težave, nakupe, komunikacijo itd. Na osnovi določenih vedenjskih vzorcev uporabnikov pa lahko podjetja tudi zaznajo goljufije ali že možnosti za njihov nastanek (Lee, 2017, str. 300).

2.5 Masovni podatki kot konkurenčna prednost

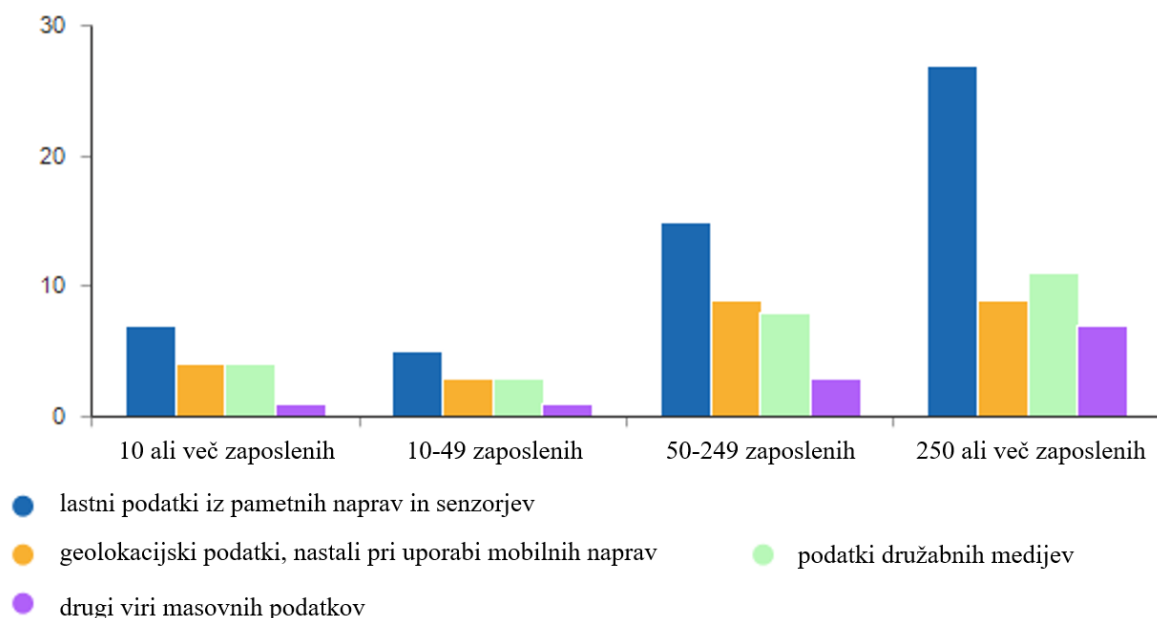
Kot navajajo Kubina, Varmus in Kubinova (2015, str. 564), predstavljata zajem in uporaba masovnih podatkov za podjetje konkurenčne prednosti in s tem omogočata podjetjem, da so korak pred konkurenco. Menijo tudi, da podjetja, ki ne izkoriščajo prednosti masovnih podatkov oziroma se ne poslužujejo uporabe masovnih podatkov, s tem na trgu dajejo prednost svojim konkurentom. Z vidika konkurenčnosti so masovni podatki lahko razumljeni kot možnost zbiranja, obdelave in analize velikih količin in obsežnih področij podatkov v določenem časovnem intervalu, ki lastniku omogočajo izdelavo poslovnih strategij in v določenih primerih tudi ustvarjanje dodatnih tržnih moči in dominantnih pozicij na trgu (Bagnoli, 2015, str. 629). Podjetja zbirajo podatke o svojih strankah na različne načine, na primer ob sklepanju pogodb v bančništvu, zavarovalništvu, nakupu avtomobila na lizing itd., podatke o svojih strankah ali prihodnjih strankah pridobivajo tudi z izvajanjem nagradnih iger, kjer sodelujoči posreduje svoje osebne podatke, z anketnimi vprašalniki, brezplačno uporabo storitev, za katero je zahtevana registracija, ipd. Prav tako tudi z zbiranjem vseh teh podatkov, ki predstavljajo veliko dodano vrednost, saj omogočajo izvajanje analiz masovnih podatkov, hkrati pa je treba upoštevati in izvajati ustrezno varnost in varovanje podatkov (Bagnoli, 2015, str. 630).

2.6 Uporaba masovnih podatkov v Sloveniji

V raziskavi Statističnega urada Republike Slovenije (v nadaljevanju SURS), ki jo je izvedel Zupan (2016b), je bilo ugotovljeno, da je leta 2015 v Sloveniji v podjetjih, kjer je bilo vsaj 10 zaposlenih, masovne podatke analiziralo 11 % podjetij. Največji odstotek podjetij, če gledamo z vidika velikosti podjetja, ki so analizirala masovne podatke, je bilo velikih podjetij, in sicer 38 %. Z vidika dejavnosti pa je bilo med podjetji, ki so analizirala masovne podatke, 14 % takih, ki se ukvarjajo s storitveno dejavnostjo, in zgolj 8 % takih, ki se ukvarjajo s proizvodno dejavnostjo (Zupan, 2016b). Največ podjetij (35 %), ki so analizirala masovne podatke, je bilo iz informacijsko-komunikacijskih dejavnosti, sledila so podjetja v dejavnosti oskrba z energijo, vodo, ravnanje z odpadki s 30 % (Zupan, 2017). Kot sem že predstavila v začetnem delu magistrskega dela, je možno zajemati podatke, ki jih uvrščamo med masovne podatke, iz različnih virov in z namenom izboljšanja lastnega notranjega poslovanja ali pa za izboljšanje zadovoljstva strank, konkurenčne prednosti ipd., kot je podrobneje predstavljeno v nadaljevanju naloge. Med podjetij, ki so v raziskavi SURS, odgovorila, da analizirajo masovne podatke, jih je največ odgovorilo, da v večini analizirajo lastne masovne podatke, ki jih zbirajo na osnovi uporabe pametnih naprav ali senzorjev

(Zupan, 2016b). Te podatke v 90 % primerov analizirajo zaposleni v podjetjih (slika 1) (Zupan, 2016b).

Slika 1: Analiza masovnih podatkov v slovenskih podjetjih za leto 2015 (v %)



Vir: Zupan (2016b).

Priložnosti uporabe masovnih podatkov pa niso le v gospodarskem sektorju, priložnosti za izboljšanje poslovanja se kažejo tudi na področju javne uprave. Tako so od aprila 2016 do februarja 2017 tudi v javni upravi Republike Slovenije (v nadaljevanju RS) izvedli pilotni projekt s ciljem raziskati možnosti uporabe sodobnih tehnologij za podatkovno analitiko in masovne podatke z namenom dviga učinkovitosti na področju javnih naročil, razvoja kadrov ter porabe virov, kjer so ugotovili tudi pozitivne učinke uporabe masovnih podatkov na aktivnosti v javnem sektorju (Kern Pipan, Bertok & Kotnik, 2017).

2.7 Primeri uporabe masovnih podatkov v globalnih podjetjih

Masovni podatki se uporabljajo v različnih dejavnostih. Nekatere, najpogostejše dejavnosti so (Özköse, Ari & Gencer, 2015, str. 1044):

- industrija avtomobilov,
- visoka tehnologija in industrija,
- nafta in plin,
- telekomunikacije,
- zdravstvo,
- trgovina na drobno in debelo,
- končni izdelki za uporabnike,

- mediji in zabava,
- potovanja in prevozi,
- finančne storitve,
- družabna omrežja in spletne storitve,
- javne storitve,
- izobraževanje in raziskovanje,
- pravo in obraba.

Številna globalna podjetja se že zavedajo pomembnosti zbiranja, hranjenja, obdelave in analize masovnih podatkov za ustvarjanje konkurenčne prednosti. Poslovne prednosti masovnih podatkov se poleg že predstavljenih kažejo predvsem v izboljšani uporabniški izkušnji, optimiziranem internem poslovanju podjetja ter ključno v izboljšanju dobička in konkurenčnosti na trgu (Alharthi, Krotov & Bowman, 2017, str. 286).

V nadaljevanju predstavljamo nekaj primerov podjetij, ki se zavedajo prednosti uporabe masovnih podatkov in jih tudi koristijo (tabela 4).

Tabela 4: Primer podjetij, ki v poslovanju uporabljajo masovne podatke

Podjetje	Vir podatkov	Tip podatkov	Namen uporabe masovnih podatkov	Vpliv na poslovanje
Etihad Airways ¹	Senzorji na letalih prek omrežja pošiljajo podatke podjetju	Stanje letala, delovanje naprav na letalu, morebitne nepravilnosti, lokacija letala.	Vzdrževanje letal, napoved morebitnih nepravilnosti oziroma t. i. napovedno vzdrževanje.	Prihranki v poslovanju podjetja
Walt Disney ²	Vstopne zapestnice, opremljene z RFID-senzorjem	Beleženje, premikanje obiskovalca in njegovo početje na posameznih lokacijah v zabaviščnem parku.	Razumevanje uporabniškega vedenja, nakupna zgodovina, čakalne dobe, preference obiskovalcev.	Izboljšanje zadovoljstva obiskovalcev zabavišnega parka in izboljšanje učinkovitosti tržnih aktivnosti. Zvišanje prihodkov na obiskovalca.

se nadaljuje

¹ Etihad Airways je letalski prevoznik, ki je bil ustanovljen leta 2003 (Etihad Airways, brez datuma). Podjetje je v lasti vlade Abu Dhabija.

² The Walt Disney Company je skupaj s svojimi hčerinskimi in pridruženimi podjetji vodilno mednarodno podjetje na področju družinske zabave in medijev z naslednjimi poslovnimi segmenti: medijska mreža, parki in kraji, studijska zabava, potrošniški izdelki in interaktivni mediji (The Walt Disney Company, brez datuma).

Tabela 4: Primer podjetij, ki v poslovanju uporabljajo masovne podatke (nad.)

Podjetje	Vir podatkov	Tip podatkov	Namen uporabe masovnih podatkov	Vpliv na poslovanje
United Parcel Services (UPS) ³	Senzorji na dostavnih vozilih	Hitrost vožnje, lokacija, stanje vozila.	Spremljanje stanja vozil in hitrosti dostave za optimiziranje dostave.	Prihranki v poslovanju podjetja.
Dublin City Council ⁴	GPS-senzorji na avtobusih.	Geolokacijski podatki vozila v realnem času.	Optimizacija voznih linij avtobusov za najboljši izkoristek.	Prihranki v poslovanju iz naslova nižje porabe goriva ter znižanja onesnaževanja ozračja.
The Daimler Group ⁵	Proizvodnja cilindrov za motorje vozil.	Več kot 500 različnih tipov podatkov, med katere sodijo tudi dimenzije in temperatura cilindra.	Napoved nastanka napak in popravek odstopanja vrednosti pred dejanskim nastankom.	Učinkovitost proizvodnje podjetja se je povečala za 25 %.

Vir: Alharthi, Krotov & Bowman (2017, str. 287).

Podjetja in institucije zbirajo masovne podatke ne glede na njihovo panogo oziroma dejavnost. Kot je razvidno iz tabele 4, med seboj panožno povsem raznolika podjetja zbirajo podatke, jih obdelujejo, analizirajo ter na njihovi osnovi tudi izboljšujejo svoje poslovanje na določenem poslovnem procesu, ki ga izvajajo. Glede na raznolikost dejavnosti podjetij tudi ta ustvarjajo različne vrste in tipe podatkov, tako na primer podjetje Etihad Airways s senzorji na letalih zbira podatke o delovanju letal z namenom, da pridobi čim več znanja o delovanju letala. Podjetje lahko pridobljene podatke uporabi kot informacije o možnosti optimizacije letal, na primer dejansko beleženje podatkov o delovanju posameznih delov letala, primerjava podatkov na dolgi rok in možnost napovedovanja potrebe po menjavi rezervnih delov, servisiranju letal ipd. Tako lahko Etihad Airways vpliva predvsem na zniževanje lastnih stroškov poslovanja oziroma optimizira svoje poslovanje.

Če podrobneje pogledamo še en primer iz tabele 4, vidimo, da podjetje Walt Disney, ki je povsem različno v primerjavi s podjetjem Etihad Airways, predvsem z vidika dejavnosti, izkorišča svoje možnosti, pridobitve in analiziranje podatkov popolnoma na drugem

³ United Parcel Services (UPS) je bilo ustanovljeno leta 1907 v ZDA kot podjetje za pošiljke (United Parcel Services, brez datuma). Podjetje je zrastle v več milijard vredno korporacijo s ciljem omogočati poslovanje po vsem svetu.

⁴ Dublin City Council ali Mestni svet Dublina je demokratično izvoljeno telo, ki je največja lokalna oblast na Irskem (Dublin City Council, brez datuma).

⁵ The Daimler Group je eden od največjih proizvajalcev prestižnih vozil in največji proizvajalec komercialnih vozil na svetu (The Daimler Group, brez datuma). Poleg proizvodnje vozil nudijo tudi financiranje, lizing, upravljanje voznega parka, zavarovanje in inovativne mobilne rešitve.

poslovnem delu podjetja, z drugačnim ciljem. Walt Disney opremlja vstopnice (zapestnice), ki jih obiskovalci nadenejo na roko, s posebnimi senzorji, ki v času obiskovalčevega obiska prenašajo podatke o obiskovalčevem gibanju v podatkovne baze podjetja. Tako lahko podjetje Walt Disney analizira in ustvarja poročila o gibanju svojih obiskovalcev po zabaviščnem parku. Za vodilne v podjetju so ti podatki pomembni, saj na njihovi osnovi pridobijo poglobljen vpogled v čakalne dobe pri posamezni atrakciji, najbolj zanimive atrakcije za obiskovalce ipd., kar vpliva na zadovoljstvo obiskovalcev.

3 MASOVNI PODATKI KOT DEL PODJETJA

Podjetja zbirajo podatke na dnevni ravni pri izvajanju poslovnih procesov oziroma pri poslovanju. Podatki nastajajo ob plačilih oziroma transakcijah, pri podpisih pogodb, kjer se podatki vnašajo v baze strank, ipd. Nekatera podjetja te podatke le hranijo na ločenih specializiranih sistemih, druga pa se odločajo za skupno hranjenje v t. i. podatkovnih jezerih. Slednja lahko na osnovi velike količine podatkov, ki se nahajajo na enotnem mestu, izvajajo analize na masovnih podatkih in tako pridobivajo poglobljene rezultate in nova spoznanja. Za zagotovitev ustreznega okolja, ki organizaciji omogoča izvajanje analiz in poročanje na osnovi masovnih podatkov, pa sta zahtevani ustreznost vzpostavitve infrastrukture in namestitve ustreznih specializiranih orodij. Podrobnejši pregled vloge in možnosti masovnih podatkov v podjetju pa so prikazane v nadaljevanju.

3.1 Vključenost masovnih podatkov v strategijo podjetja

Vedno več organizacij, predvsem korporacij, se odloča za vodenje na osnovi podatkovne informiranosti z namenom ustvarjanja bolj usmerjenih storitev, zniževanja tveganj in izboljšanja učinkovitosti poslovanja podjetja (Fan, Han & Liu, 2014, str. 296). Organizacije nameščajo specializirane programe podatkovne analitike z namenom zbiranja, hranjenja, obdelave in analize velikih podatkovnih nizov iz nabora različnih virov. Ključni namen uporabe omenjenih programov je identificiranje poslovnega stanja in poslovnih priložnosti za omogočanje podpore poslovnemu odločanju (Fan, Han & Liu, 2014, str. 296).

Masovni podatki so prisotni na različnih področjih v različnih industrijah oziroma dejavnostih poslovanja. Prisotni so v proizvodnji, logistiki, na področju finančnih storitev, zdravstva, vlade, tehnologije itd. Strokovnjaki, ki so dejavni na področju informacijske tehnologije, govorijo o masovnih podatkih ter njihovem vplivu in priložnostih v poslovanju. Veliko jih verjame, da bodo masovni podatki ustvarili nove organizacijske priložnosti in vrednosti podjetja (Davenport, Barth & Bean, 2012, str. 22).

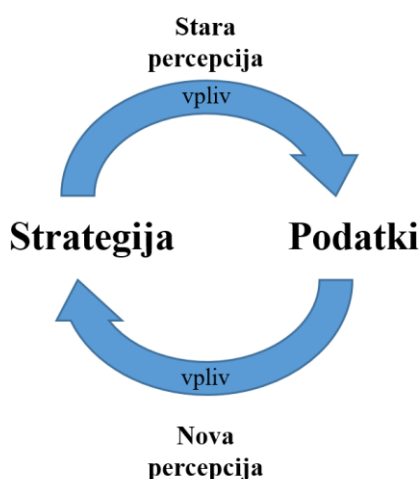
Izvršni delavci v različnih podjetjih in organizacijah raznovrstnih industrijskih dejavnosti si prizadevajo za več znanja na področju masovnih podatkov; iščejo usmeritve, dobre prakse in boljše razumevanje področja. Predvsem se ukvarjajo z vplivom masovnih podatkov na poslovanje in procese organizacij ter z vidika vpliva masovnih podatkov na sprejemanje

strateških odločitev (Mazzei & Noble, 2017, str. 406). Odločevalci konstantno iščejo in pričakujejo dodatne informacije za bolj informirano odločanje o poslovanju, hkrati morajo biti tudi zmožni razumeti masovne podatke, da lahko podprejo nadaljnji razvoj tradicionalnega načina sprejemanja odločitev (Elgendy & Elragal, 2016, str. 1073).

Kot navajata Mazzei in Noble (2017, str. 406), sta bila v preteklosti področje masovnih podatkov ter njihova vloga v upravljanju podjetja nekoliko zanemarjena oziroma pomanjkljivo obravnavana v akademski literaturi. Vse večje količine podatkov in napredne analitične metode pa zahtevajo ponovno proučitev predpostavk o načinu poslovanja ter o vplivu na preoblikovanje organizacije in podjetja (Mazzei & Noble, 2017, str. 407). Z vse večjo količino ustvarjenih podatkov digitalizacija omogoča spremembe poslovanja in s tem tudi vpliv na strateški del organizacije (Mazzei & Noble, 2017, str. 406).

Najbolj uspešna podjetja uporabljajo analitiko petkrat pogosteje kot podjetja, ki so manj uspešna (LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011, str. 23). Kot navajata Mazzei in Noble (2017, str. 407), je v današnjem času treba oblikovati in definirati strategijo podjetja na osnovi zbranih in analiziranih podatkov, saj lahko imajo ti precejšen vpliv na korporativno strategijo podjetja. Kot je razvidno iz slike 2, po Mazzeiu in Noblu (2017) obstaja odnos med strategijo in podatki, ki so med seboj močno povezani. Obdobje, kjer so podjetja oblikovala potrebo po določenih metrikah in zbirala pripadajoče podatke na osnovi določene korporativne strategije, v današnjem času postaja nepredstavljivo, saj danes podatki nastajajo na različnih ravneh podjetja in izziv, s katerim se podjetja soočajo, je, kako iz zbranih podatkov »izluščiti« dodano vrednost, s katero bodo vplivali na boljše rezultate poslovanja.

Slika 2: Razvoj odnosa med podatki in strategijo



Vir: Mazzei & Noble (2017, str. 407).

3.2 Prihodnost masovnih podatkov v podjetjih

Na svetovni ravni je do leta 2020 pričakovana 40 % letna rast količine podatkov, kot predstavlja de Waal-Montgomery (2015). Kot zgled, kako hitro narašča količina ustvarjenih podatkov v zadnjih letih, dodajam prikaz rasti podatkov skozi leta od leta 2005 do leta 2015 (slika 3).

Slika 3: Pregled rasti količine podatkov skozi leta (v zettabajtih)



Vir: Yaqoob in drugi (2016, str. 1234).

Vedno večje količine podatkov nastajajo z vse večjo uporabo naprav, ne zgolj računalnikov in pametnih telefonov, temveč tudi drugih naprav, ki so povezane v internet in jih uporabljamo tako končni uporabniki v priložne namene kot tudi na dnevni ravni podjetja in ustanove, ki potrebujejo internet za dnevno delo.

Na začetku najprej pogledajmo razliko med t. i. internetom računalnikov, kot so ga poimenovali Qin in drugi (2016, str. 137), ter internetom stvari. Internet računalnikov se primarno ukvarja oziroma je usmerjen v naprave, kot so računalniki, internet stvari pa ni zgolj usmerjen v računalnike, temveč na širši nabor naprav, kot so ure, gospodinjske naprave, čevlji, oblačila, avtomobili ipd., ki so povezani z internetom (Qin in drugi, 2016, str. 137).

Internet stvari lahko imenujemo tudi internet vsega ali industrijski internet, saj gre za globalno mrežo povezanih naprav, ki so zmožne medsebojne komunikacije (Lee & Lee, 2015, str. 431). Največja in prava vrednost interneta stvari za podjetja je, kadar so naprave med seboj povezane, lahko med seboj izmenjujejo podatke in so zmožne izmenjave podatkov s sistemi zunanjih izvajalcev, sistemi za pomoč uporabnikom, poslovno-

inteligentnimi aplikacijami in poslovno analitiko (Lee & Lee, 2015, str. 431). Kot navajajo Qin in drugi (2016, str. 138), je internet stvari eden od glavnih dejavnikov, ki ustvarjajo potrebo po analitiki na področju masovnih podatkov. Vedno večje količine podatkov, ki nastajajo z uporabo naprav interneta stvari namreč ustvarjajo masovne podatke (Malek in drugi, 2017, str. 430).

Ključne tehnologije interneta stvari, ki ustvarjajo masovne podatke, so (Lee & Lee, 2015, str. 432):

- RFID – identifikacija radijske frekvence,
- WSN – brezžično omrežje senzorjev,
- Middleware – programska oprema, ki služi kot most oziroma povezava med operativnimi sistemi, podatkovnimi bazami in aplikacijami na omrežju,
- računalništvo v oblaku (ang. Cloud computing),
- aplikativna programska oprema interneta stvari.

S hitrim dostopom do podatkov in informacij internet stvari omogoča ustvarjanje visoko inovativnih in učinkovitih storitev, ki omogočajo definiranje različnih poslovnih modelov podjetij (Metallo, Agrifoglio, Schiavone & Mueller, 2018).

Obstaja tudi drugačen pogled glede obstoječih podatkov, ki jih že ustvarjajo naprave interneta stvari. Takšen pogled ima denimo podjetje HP, ki meni, da podatki, ki so že ustvarjeni na račun naprav interneta stvari, še ne sodijo med masovne podatke. Podjetje HP je napovedalo, kot navajajo Yaqoob in drugi (2016, str. 1234), da bo do konec leta 2030 število senzorjev na svetu preseglo 1 trilijon, in po njihovem mnenju bodo ti podatki, ki jih ustvarja internet stvari, takrat postali del masovnih podatkov.

3.3 Masovni podatki in analitika masovnih podatkov kot del prihodnosti podjetij

Leta 2011 se je pojavil koncept t. i. industrija 4.0 (ang. Industry 4.0), ki je vseboval ključne tehnološke inovacije na področju izvedbe procesov avtomatizacije, kontrole in nadzora ter informacijske tehnologije (Santos in drugi, 2017, str. 751). Industrija 4.0 tako z vidika industrijske revolucije predstavlja 4. industrijsko revolucijo, kjer je bila 1. industrijska revolucija leta 1780 z izumom parnega motorja in mehaničnih statev, 2. industrijska revolucija je bila okrog leta 1870 s pričetkom uporabe električnih motorjev in naftnega goriva. Sledila je še 3. industrijska revolucija, ki naj bi bila okrog leta 1970 in temelji na uporabi računalniško vodenih sistemov in robotov v industrijski proizvodnji (Santos in drugi, 2017, str. 751). Za razvoj industrije 4.0 pa je potrebnih 6 pogojev (Santos in drugi, 2017, str. 751):

- notranje sodelovanje in komuniciranje informacij med sistemi, ljudmi – zlitje fizičnega in virtualnega sveta,
- izvajanje aktivnosti v realnem času – sprejemanje odločitev v realnem času,

- virtualizacija – pametne proizvodnje, spremljanje in nadzor izvajanja procesov na osnovi številnih senzorjev, nameščenih po proizvodnji oziroma območju podjetja,
- decentralizacija – postopoma se bo pričela brisati meja med podajanjem navodila oziroma določanjem nalog napravam ter omogočanjem, da bodo tudi naprave lahko na osnovi zajetih podatkov posredovale povratne informacije o izvedbi procesov, predlagale izboljšave ipd.,
- usmerjenost v storitve – združevanje uporabe storitveno usmerjene programske opreme in storitev ter izdelkov internet stvari,
- modularnost – procesi izvedbe so v skladu s povpraševanjem in zahtevami, združevanje in razdruževanje modulov in izvedbe, z zagotovitvijo fleksibilnosti, možnosti za spremembo izvedbenih aktivnosti naprav.

Na osnovi predstavljenih pogojev za razvoj industrije 4.0 pa Santos in drugi (2017) predstavljajo 6 ključnih tehnologij, kamor sodijo: varnost na spletu, internet stvari, masovni podatki, mobilna in virtualna realnost, spletni oblak ter proizvodnja (slika 4). Te tehnologije ustvarjajo vedno večje količine podatkov.

Slika 4: Tehnologije industrije 4.0

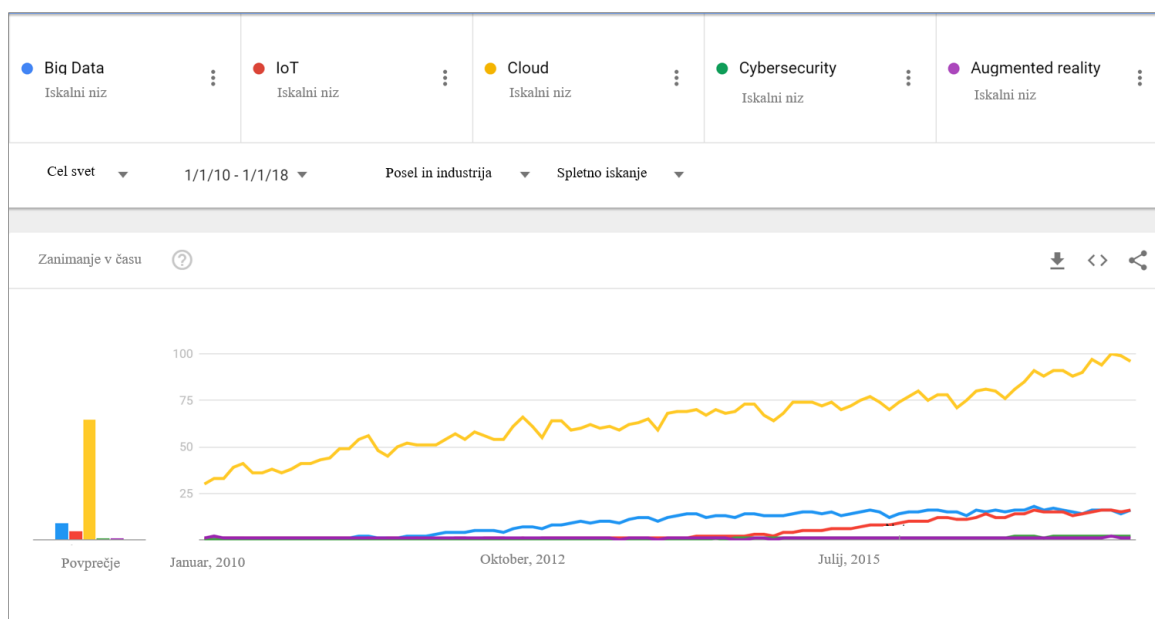


Vir: Santos in drugi (2017, str. 751).

Glede na podatke od januarja 2010 do januarja 2018 o spletnem iskanju po ključnih besedah: Big Data (slov. masovni podatki), IoT (slov. internet stvari), Cloud (slov. oblak), Cybersecurity (slov. varnost na spletu) in Augmented reality (slov. virtualna resničnost), ki so bile kategorizirane med iskalne ključne besede, povezane s poslovnim področjem na globalni ravni, lahko opazimo, da ima najvišji trend iskanj, ki je še vedno rastoč, ključna beseda »Cloud«, in sicer je dosegla najvišji možen interes skozi čas, z vrednostjo tudi do

100⁶. Nato sledijo v zaporedju z vidika priljubljenosti: »Big Data«, »IoT«, »Cybersecurity« in »Augmented reality« (slika 5). Kot je razvidno iz slike 5, je trend pričetka rasti iskanja ključnih besed, povezanih z »Big Data«, možno opaziti nekje od začetka leta 2012, kjer se je rast nato nekje v letu 2015 stabilizirala in ohranjala raven interesa vse do leta 2018. Trend rasti ključne besede »IoT« pa vidimo, da se prične v začetku leta 2015 zviševati, kjer tudi v drugi polovici leta 2017 dohiti stopnjo interesa ključne besede »Big Data«.

Slika 5: Spletno iskanje na globalni ravni, po ključnih besedah v angleškem jeziku od 2010 do 2018

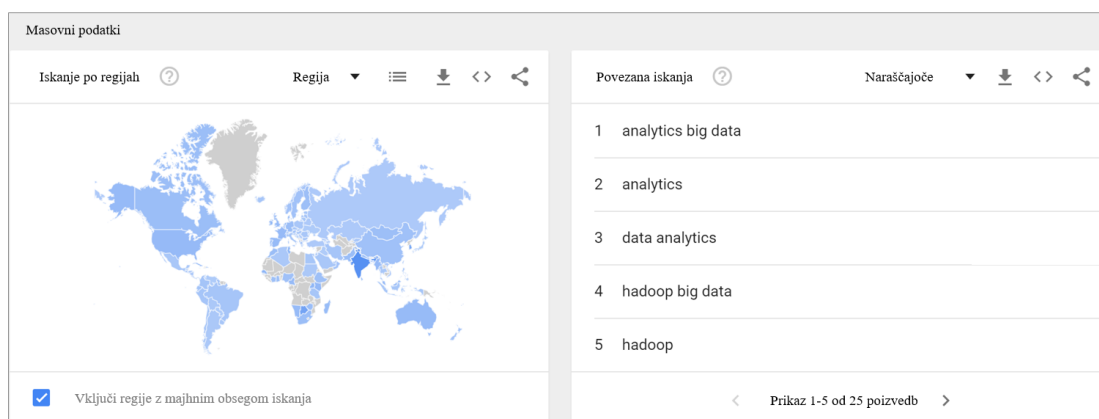


Vir: Google Trends (2018).

Če se osredotočimo naprej na masovne podatke, lahko vidimo, da se pretežno iščejo ključne besede, ki so povezane z analitiko oziroma izvajanjem analiz na zbranih podatkih, ter da gre za področje zanimanja, ki je zanimivo skoraj na vseh celinah sveta (izjemi sta Afrika in Grenlandija) (slika 6).

⁶ Vrednost predstavlja relativni interes iskalne ključne besede v primerjavi z najvišjo točko grafa za določeno področje v določenem časovnem okviru. Vrednost na primer 100 predstavlja najvišjo stopnjo priljubljenosti za iskano ključno besedo.

Slika 6: Iskane sorodne ključne besede, povezane s ključno besedo »Big Data«



Vir: Google Trends (2018).

Saggi in Jain navajata (2018), da lahko sestavimo popis tehnologij, ki sodijo v sklop tehnologij prihodnosti. Te tehnologije so določene z vidika razvijajočih se raziskovalnih tehnologij, kjer določene že, določene pa še bodo pokrivalo posebna področja znanstvenega zanimanja in s katerimi bo omogočeno raziskovanje specifičnih področij (Saggi & Jain, 2018). Kot je razvidno iz tabele 5, sodi tudi analitika masovnih podatkov med tehnologije prihodnosti, katerih namen oziroma zanimanje je osredotočeno predvsem na pametna parkirišča, pametna mesta, internet stvari ipd. (tabela 5) (Saggi & Jain, 2018).

Tabela 5: Seznam razvijajočih se raziskovalnih tehnik in njihova področja zanimanja

Tehnike prihodnosti	Interesna področja	Področje zanimanja
Računalništvo v oblaku (ang. Cloud computing)	– virtualizacija – programsko določeno omrežje – upravljanje virov	– masovni podatki kot storitev – analitika podatkov kot storitev
Umetna inteligenca (ang. Artificial Intelligence)	– optimizacija – nevronske mreže – rudarjenje masovnih podatkov – rudarjenje interneta stvari	– pomoč pri razvoju inteligentnih naprav
Računalništvo biološkega navdiha (ang. Bio-inspired Computing)	– imunski sistem – sistem Linder Mayer (L-sistem)⁷ – membransko računalništvo	– drone⁸ za nadzor insektov

se nadaljuje

⁷ L-sistem je en izmed modelov, ki so bili razviti z namenom proučevanja življenjskih navad in karakteristik populacije umetnega življenja (ang. artificial life population) (Feng & Chen, 2017).

⁸ Drone je v tehnološkem smislu zračno plovilo oziroma zračno vozilo brez posadke. Gre za letečega robota (TechTarget IoT Agenda, 2016).

*Tabela 5: Seznam razvijajočih se raziskovalnih tehnik in njihova področja zanimanja
(nad.)*

Tehnike prihodnosti	Interesna področja	Področje zanimanja
Kvantno računalništvo (ang. Quantum Computing)	– elektronsko računalništvo – optično računalništvo – kvantna ura	– kvantni računski modeli – kvantni model strojnega učenja – dinamično kvantno združevanje – izboljšana optimizacija – boljša najdljivost – povečana varnost – prepoznavanje vzorcev, kemično modeliranje
Kvantna kriptografija (ang. Quantum cryptography)	– enkripcija javnega ključa – shema podpisa	– pomaga pri izvajanju kriptografskih nalog – obdelava slik – napredna varnost
Robno računalništvo (ang. Edge computing)	– megleno računalništvo – mobilno robno računalništvo – oblak	– z zmanjšanjem računanja uporablja uporabnike tako, da računanje spusti do roba omrežja
Analitika masovnih podatkov in strojno učenj (ang. Big Data Analytics & Machine Learning)	– razvrščanje in združevanje – napovedno modeliranje – analitika v realnem času – poglobljeno učenje	– finance in računovodstvo – internet stvari v agrikulturi – pametni transporti – pametno zdravstvo – pametna mesta – pametna mreža – pametni sistem inventarja – pametno parkiranje

Vir: Saggi & Jain (2018).

4 KVANTITATIVNA RAZISKAVA UPORABE MASOVNIH PODATKOV

V tem poglavju je predstavljena uporabljena metodologija empiričnega dela naloge, kjer sta tudi podrobno zapisana in predstavljena način pridobitve informacij ter način sestave in izvedbe spletnega vprašalnika. Po zaključku vprašalnika je sledila obdelava podatkov, ki je tudi predstavljena v prvem delu tega poglavja. Nato sledita še predstavitev namena in ciljev naloge ter v zadnjem sklopu tega poglavja analiza pridobljenih in obdelanih podatkov iz spletnega vprašalnika.

4.1 Metodologija

Empirični del naloge temelji na proučitvi ozaveščenosti in pomenu uporabe masovnih podatkov v velikih slovenskih podjetjih. Raziskavo sem izvedla s pomočjo spletnega vprašalnika, ki je temeljil na sekundarno zbranih podatkih znanstvenih in strokovnih avtorjev na globalni ravni o že znanih načinih uporabe, o priložnostih in o nevarnostih masovnih podatkov.

Vprašalnik sestavlja 19 vprašanj, ki temeljijo na obstoječih znanstvenih in strokovnih virih. Prvi sklop se nanaša na poznavanje, uporabo, vpletenost vodstva, znanje ter strateško vrednost masovnih podatkov za podjetje. Začetna vprašanja spletne ankete respondente razdelijo med tiste, ki poznajo ali so že slišali za izraz »masovni podatki« ter jih tudi uporabljajo v podjetju, ter tiste, ki izraza sploh ne poznajo. Pri zasnovi vprašalnika sem izhajala iz sklepa, da če respondent izraza ne pozna oziroma se v podjetju ne ukvarjajo z zbiranjem in obdelovanjem podatkov, takšno podjetje sodi med tista, ki jih ni smiselno vključevati v nadaljevanju oziroma jih lahko kasneje v analizi uvrstimo med podjetja, ki bi lahko predstavljala potencial za izobraževanje s področja masovnih podatkov ali pa v njihovi panogi poslovanja masovni podatki ne predstavljajo dodane vrednosti za njihovo poslovanje.

V začetnem delu vprašalnika, torej pri prvem vprašanju Q1: »Ali poznate oziroma ste že slišali za masovne podatke?« preverjam poznavanje, sprašujem zgolj z namenom preprostejše razporeditve in zasnove vprašanj v nadaljevanju vprašalnika. Pri vprašanju Q2, ki se pojavi respondentom, ki so na vprašanje Q1 odgovorili z »ne«, se naslonim na definicijo masovnih podatkov po Alharthiju, Krotovu in Bowmanu (2017, str. 286) z namenom, da pojasnim pomen tematike, po definiciji pa sledi vprašanje, ali bi po seznanitvi pričeli razmišljati o uvedbi uporabe masovnih podatkov v njihovem podjetju. Respondente, ki so pri vprašanju Q1 odgovorili z »da« oziroma da izraz poznajo, pa sem pri vprašanju Q3 vprašala, ali v njihovem podjetju obdelujejo podatke. V nadaljevanju vprašalnika se vprašanja vse do sklopa o demografskih podatkih nanašajo na respondente, ki izraz poznajo.

Pri vprašanju Q4 sem respondente vprašala, iz katerih virov zbirajo podatke v njihovem podjetju. Kot odgovor sem vnaprej ponudila možnosti, kjer sem prav tako kot pri Q2 izhajala iz virov, ki jih omenjajo tudi Alharthiju, Krotovu in Bowmanu (2017, str. 288). Pri vprašanju Q5 sem preverjala metode oziroma tehnike, ki jih v podjetjih uporabljajo, kjer pa sem se osredotočila predvsem na analitični tehniki tekstovnega rudarjenja in napovedno analitiko. Kot teoretično znanstveno izhodišče je bil uporabljen zapis, ki se v delu osredotoča predvsem na omenjeni metodi obdelave in analize podatkov ter termina masovnih podatkov (Finlay, 2014, str. 104 in 179).

Pri nadaljnjih vprašanjih od Q6 do Q12 bi lahko vprašanja uvrstili v sklop, ki se nanaša na notranje poslovanje podjetja. Masovni podatki prikazujejo vpliv in pomembnost aktivnosti notranjega poslovanja na stroškovno učinkovitost, hitrost izvedbe, kakovost proizvodnje ipd. v podjetju.

Kot navajata McAfee in Brynjolfsson (2012), se mora podjetje, ki je vodeno na osnovi podatkov (ang. data-driven), vprašati »Kaj vemo?« in ne »Kaj mislimo?«. Ta in podobna vprašanja, ki določajo tudi kulturo podjetja, kot navajata McAfee in Brynjolfsson (2012), nakazujejo podporo vodstva in interes za vključevanje v širšo sliko podjetja, kamor sodi tudi strategija podjetja. Na začetku, pri vprašanju Q6, respondente sprašujem o vključenosti masovnih podatkov in z njimi povezanih aktivnosti v strategijo podjetja, kot predstavljata pomembnost »vpetosti« masovnih podatkov v strategijo podjetja tudi Mazzei in Noble (2017, str. 406–407). Zanimalo me je, ali merijo tudi učinek uporabe masovnih podatkov na rezultate podjetja, zato pri vprašanju Q7, kjer sem izhajala iz izvedene raziskave Davenporta in Beana (2017, str. 3), preverim tudi to. Pri respondentih, ki so na vprašanje Q7 odgovorili z »da«, sem pri vprašanju Q8 preverila, kak učinek opažajo, po zgledu Davenporta in Dyché (2013, str. 22), ki omenjata, da se je treba zavedati, da so lahko učinki uporabe masovnih podatkov pozitivni ali negativni. Z vidika finančnih vlaganj v razvoj in upravljanje masovnih podatkov so odgovarjali respondenti pri vprašanjih Q9 in Q10, kjer sem pri vprašanju Q9 spraševala o višini vlaganj v zadnjih 5 letih po zgledu vprašalnika, ki sta ga izvedla Davenport in Bean (2017, str. 7), ter pri vprašanju Q10 o področju, za katero so bila ta vlaganja porabljena. Davenport in Dyché (2013, str. 9–11) navajata različna področja, ki zahtevajo finančne vloške in razvoj, kot so vizualizacija podatkov, zaledni sistemi itd. Kot navaja Lee (2017, str. 299), je z aktivnostmi na področju masovnih podatkov možno vplivati na različna področja v podjetju. Primeri področij, na katera je možno vplivati, so zniževanje operativnih stroškov, vzpostavitev novih poti, monetiziranje masovnih podatkov, večja konkurenčnost ipd., kot navajata tudi Davenport in Bean (2017, str. 8). Izhajajoč iz teh spoznanj, sem sestavila vprašanje Q11, kjer sem respondente vprašala, na katera področja v njihovem podjetju oni vplivajo z aktivnostmi. S kadrovskega vidika ter vodstvene podpore sem, izhajajoč iz zapisov Braganza, Brooksa, Nepelskega, Alija in Moroa (2017, str. 335), pri vprašanju Q12 vprašala glede naklonjenosti in podpore vodstva pri aktivnostih, povezanih z masovnimi podatki, ter pri vprašanju Q13, ali se soočajo z izzivi zadostnega števila kadra z ustreznim znanjem s proučevanega področja (Alharthi, Krotov & Bowman, 2017, str. 288). Za konec, pred prehodom na sklop demografskih vprašanj, sem respondente vprašala še o izzivih, s katerimi se soočajo, ter tudi, ali morda razmišljajo o prihodnosti in vedno hitreje prihajajočih izdelkih in storitvah t. i. interneta stvari, ki bodo generirale še več podatkov. Leejeva navaja (2017, str. 301-302) nekaj možnih izzivov, s katerimi se podjetje, ki obdeluje, analizira in posluje z masovnimi podatki, lahko sreča v času izvedbe aktivnosti, zato sem tudi sama pri vprašanju Q14 spraševala o izzivih, s katerimi se srečujejo v velikih slovenskih podjetjih.

Z mislijo na prihodnost sem pri vprašanju Q15 na osnovi Krotovih (2017, str. 1-2, 11) zapisov o prihajajočem internetu stvari in s tem povezano rastjo količine podatkov, ki bodo predstavljala še dodaten izziv, a hkrati tudi še dodatne poslovne priložnosti, respondente vprašala, ali se bodo v prihodnosti tudi oni posluževali uporabe napovedne analitike na področju interneta stvari.

V zadnjem delu vprašalnika so sledila demografska vprašanja o dejavnosti podjetja respondenta (Q16), položaju respondenta v podjetju (Q17) in spolu respondenta (Q18).

Pri vprašanju Q16: »Izberite panogo, v katero sodi vaše podjetje« se odgovori v analizi razlikujejo od prvotnih, ki so bili zabeleženi v anketiranju. Pri obdelavi dobljenih podatkov, sem iz zadnjega ponujenega odgovora »Drugo:«, kjer so respondenti imeli možnosti sami zapisati odgovor, združila navedene podatke v smiselne sklope z namenom lažje interpretacije in analize rezultatov. Tako sem združila obstoječo skupino »Gostinstvo« s turizmom, ki so ga navedli respondenti v novi skupini, ki sem jo poimenovala »Gostinstvo in turizem«, ter obstoječo skupino »Energetika« in s strani respondentov navedeno regulacijo ogrevanja, prezračevanja in klimatizacije v »Energetika in regulacija ogrevanja, prezračevanja in klimatizacije«. Respondenti so navedli tudi nekaj novih sklopov, tako sem tudi te združila v smiselne skupine in na novo dodala naslednje skupine: »Avto industrija«, »Proizvodnja«, »IT in informatika« ter »Kemična industrija«. Skladno z združitvijo in novo ustvarjenimi skupinami sem tako tudi ustrezno prenesla in uredila vrednosti oziroma pripadajoče frekvence posameznih odgovorov. Pri vprašanju Q16, nisem respondentom ponudila v naprej podanih odgovorov v skladu s standardno klasifikacijo dejavnosti 2008, V2 (v nadaljevanju SKD_2008, V2) (Braunsberger, Hlavaty, Schlamberger & Stevanovič, 2010), temveč sem sama sestavila nabor odgovorov na podlagi analize, ki jo je izvedlo podjetje Intel Corporation (2012) ter nekaj ostalih predhodno omenjenih vprašalnikov. Za nadaljnja raziskovanja pa je vsekakor priporočljivo upoštevati SKD_2008, V2, ki se uporablja za razvrščanje gospodarskih dejavnosti.

Z namenom lažje interpretacije sem tudi pri vprašanju Q17, kjer so respondenti označili ali dopisali svoj položaj v podjetju, združila dobljene odgovore v tri skupine: »Uprava«, »Vodja« in »Izvedbena funkcija«. Respondenti so največkrat izbrali odgovor »Drugo:« in odgovorov zapisali kar sami (16 zapisov). Glede na podane odgovore je bilo možno ustvariti tri skupine, saj so odgovori sovpadali z vidika položaja respondenta, ne pa tudi področja dela, z že določenimi skupinami. Skupina »Uprava« je sestavljena zgolj iz odgovorov, označenih v vprašalniku, brez dodatnih zapisov, skupina »Vodja« je sestavljena iz izbora med podanimi odgovori ter tudi združitve s podanimi dopisi, ki jih je bilo dodatno še 6. Še zadnja skupina, ki sem jo poimenovala »Izvedbena funkcija«, pa je sestavljena zgolj iz dopisov respondentov, in sicer vseh 10.

Spletno anketiranje sem izvedla s pomočjo orodja 1KA. Populacijo proučevanja so predstavljala velika slovenska podjetja. Velika podjetja so po Zakonu o gospodarskih družbah (ZGD-1), Ur. l. RS, št. 42/2006 definirana kot podjetja, ki morajo izpolnjevati vsaj dva od treh kriterijev: več kot 250 zaposlenih, čisti prihodki od prodaje presegajo 40.000.000 EUR in vrednost aktive presega 20.000.000 EUR (ZGD-1, 55. člen). Po izvedeni analizi s pomočjo orodja bizi.si sem ugotovila, da je v Sloveniji 253 velikih podjetij. Pri zasnovi dispozicije sem sprejela odločitev, da bom za analizo rezultatov uporabila vzorec 50 podjetij. Ocenila sem, da je velikost vzorca zadostna za sklepanje zaključkov o populaciji.

Spletno anketo sem ustvarila 28. septembra 2017 ob 9.30 in zagnala 3. oktobra 2017. Po zagonu ankete sem v naslednjih dneh (od 3. oktobra do 9. oktobra 2017) poslala na bazo 238 elektronskih naslovov velikih slovenskih podjetij od skupno 253 zajetih podjetij v populacijo, saj je bilo nekaj elektronskih naslovov neveljavnih, niso obstajali ali jih ni bilo možno pridobiti. V elektronsko sporočilo sem vključila povezavo do anketnega vprašalnika. Poleg povezave do vprašalnika sem sporočilu dodala tudi nagovor, kjer sem na kratko predstavila sebe ter cilj in namen ankete, ki so ga prejeli. Anketa je sestavljena iz 19 vprašanj na 18 straneh, trajanje izpolnjevanja je bilo predvideno na 3 minute in 9 sekund, vendar je bilo povprečno trajanje izpolnjevanja 1 minuto in 31 sekund. Na povezavo do spletne ankete v elektronskem sporočilu je kliknilo 72 prejemnikov sporočila.

Spletno anketiranje sem zaključila 12. novembra 2017 ob 17.49, saj sem zadnji odgovor zabeležila 24. oktobra 2017 in tako ocenila, da bo vzorec pridobljenih odgovorov zadostoval. Po premisleku sem tudi ugotovila, da glede na anonimno izvedeno anketiranje ne vem, kdo od vabljenih k izpolnitvi je dejansko izpolnil anketo, tako tudi nisem imela možnosti ponovno povabiti ali spomniti zgolj tistih, ki se niso odzvali na prejeto sporočilo. Povabilo k izpolnitvi ponovno vsem, tudi tistim, ki so anketo že izpolnili, pa se mi je zdelo preveč invazivno, zato se nisem odločila za ponovno pošiljanje vabil. Razmišljala sem sicer tudi o drugih možnostih, na primer objavi povezave na družabnih omrežjih, kot je na primer LinkedIn, ali objavi oglasa prek Google AdWords, vendar bi s takšno objavo dejansko uničila svojo ciljno skupino, ki sem jo želela proučiti, saj bi se tako anketiranja lahko udeležil vsakdo, tudi fizične osebe, ki na primer sploh niso zaposlene v podjetjih.

4.2 Namen in cilji raziskave

Primarni namen naloge je bil proučiti ozaveščenost, obstoječe stanje in naklonjenost uporabi masovnih podatkov v velikih slovenskih podjetjih. Masovni podatki predstavljajo priložnosti, kot so: lažje poslovno odločanje na osnovi podatkov, ustvarjanje konkurenčne prednosti, boljše poznavanje strank, predvidevanje njihovih nadaljnjih nakupnih odločitev, možnost priprave prilagojene ponudbe posamezniku ipd. Namen naloge je bil tudi prikazati nabor poslovnih priložnosti, ki izhajajo iz uporabe masovnih podatkov in bi lahko bile tudi zanimive za velika podjetja v Sloveniji.

Temeljni cilj magistrskega dela je preveriti ali obstajajo v velikih slovenskih podjetjih na področju masovnih podatkov neizkoriščene poslovne priložnosti. Poslovne priložnosti se kažejo v različnih oblikah, ena izmed njih je zagotovo konkurenčna prednost. Konkurenčno prednost lahko gradi podjetje v smislu razlikovanja ponudbe, ki je lahko prilagojena posameznemu obstoječemu ali potencialnemu kupcu. Podjetje, ki v svojem poslovanju uporablja masovne podatke, lahko končnemu uporabniku ponudi prilagojeno ponudbo, še preden ta dejansko začuti potrebo po izdelku. Tako lahko podjetje prehití konkurenta s pravočasnim nagovorom stranke in prilagojeno ponudbo, hkrati mu tak način poslovanja omogoča tudi optimizacijo internih procesov oziroma zniževanje stroškov.

4.3 Raziskovalna vprašanja

Raziskovalna vprašanja bodo usmerjena k ugotavljanju poznavanja, potreb in zmožnosti uporabe masovnih podatkov velikih slovenskih podjetij.

Raziskovalno vprašanje:

Kako dobro v velikih slovenskih podjetjih izkoriščajo možnosti za izboljšanje, prilagajanje ali širitev poslovanja podjetja na osnovi uporabe masovnih podatkov?

Eksperimentalne hipoteze:

H1: Večina velikih podjetij v Sloveniji se ne poslužuje naprednih statističnih analiz masovnih podatkov, ki vključujejo tehnike napovedne analitike in podatkovnega rudarjenja.

H2: Večina velikih slovenskih podjetij si želi več analitičnih podatkov, ki bi jim pomagali izboljšati poslovanje.

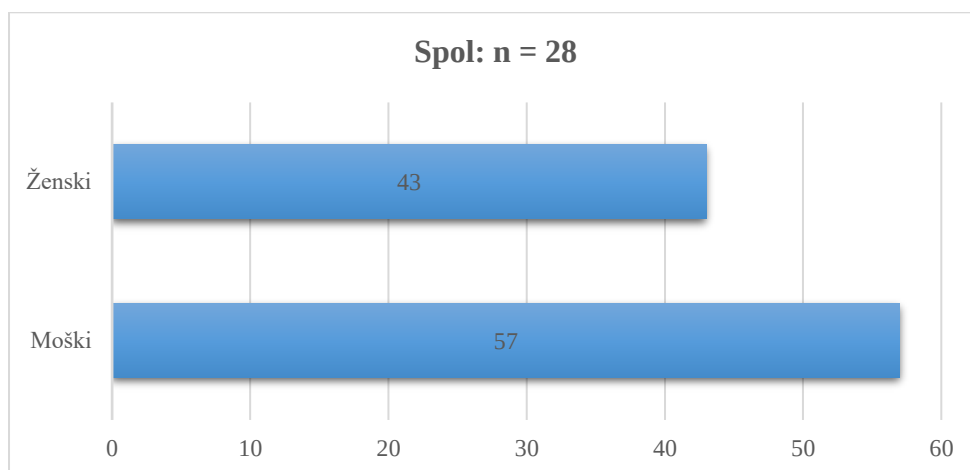
4.4 Analiza

Pridobljene podatke sem obdelala na različne načine z različnimi orodji in metodami, kar sem predstavila v poglavju 4.1, kjer sem predstavila metodologijo izvedbe raziskave. V nadaljevanju naloge so predstavljeni rezultati raziskave. Rezultati so prikazani v opisni, grafični in tabelarični obliki.

Kot sem že zapisala v poglavju 4.2, kjer sta podrobno predstavljena namen in cilj naloge, je cilj naloge preveriti, ali obstajajo neizkoriščene poslovne priložnosti na področju masovnih podatkov v velikih slovenskih podjetjih. Tako sem si v nalogi zastavila tudi vprašanje, kako dobro v velikih slovenskih podjetjih izkoriščajo možnosti za izboljšanje, prilagajanje ali širitev poslovanja podjetja na osnovi uporabe masovnih podatkov.

Na začetku analize sledi pregled rezultatov demografskih vprašanj, na osnovi katerih bo lažje razumeti, na kateri spol, položaj v podjetju ter dejavnost podjetja se v nadaljevanju nanašajo podatki analitičnega pregleda in ugotovitev. K udeležbi v spletnem anketiranju je bilo vabljenih 238 respondentov oziroma je bilo elektronsko sporočilo z vabilom poslano na 238 elektronskih naslovov velikih slovenskih podjetij, kjer jih je sodelovalo 72, vendar je zgolj 36 respondentov odgovorilo na vprašanja na način, da odgovori štejejo kot veljavni. Tako sem analizo, ki sledi v nadaljevanju, naredila na osnovi veljavno izpolnjenih vprašanj. Med respondenti je odgovarjalo 57 % moških in 43 % žensk oziroma 16 moških in 12 žensk (slika 7).

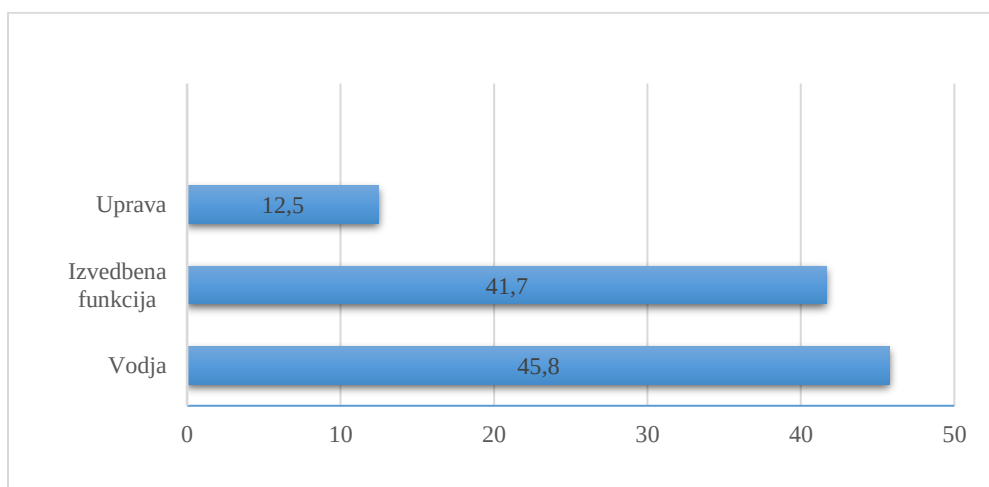
Slika 7: Odstotek respondentov glede na spol



Vir: lastno delo.

Z namenom, da pridobim grob vpogled v organizacijsko strukturo podjetja oziroma položaj respondentov, ki se ukvarjajo s področjem masovnih podatkov, sem v elektronskem sporočilu, kjer se je nahajalo vabilo k izpolnitvi spletne ankete, predlagala, da spletno anketo izpolni nekdo v podjetju, ki se spozna, je večš oziroma dela na področju poslovne inteligence, obdelave podatkov ali informatike ipd. Kot sem ugotovila z vprašanjem Q17, kjer so respondenti označili svoj položaj v podjetju, je večina vprašanih odgovorila, da so vodje določenega oddelka (45,8 %) ali pa izvajajo operativno funkcijo (41,7 %). Nekoliko presenetilo pa je tudi dejstvo, da so od skupno 24 veljavnih odgovorov na spletno anketo odgovorili tudi 3 člani uprave, kjer je sodeloval 1 član uprave za podatke (ang. Chief Data Officer) in 2 člana uprave za trženje (ang. Chief Marketing Officer), kar predstavlja 12,5 % (slike 8).

Slika 8: Odstotek respondentov na določenem položaju v podjetju: uprava, vodja in izvedbena funkcija



Vir: lastno delo.

Kot je prikazano v tabeli 6, večina sodelujočih podjetij oziroma respondentov, ki so zaposleni v podjetjih, ki sem jih povabila k izpolnitvi spletne ankete, izhaja iz dejavnosti logistika (5 respondentov je navedlo dejavnost logistika) in proizvodnje dejavnosti (5 respondentov je navedlo panogo proizvodnje). Na drugem mestu sledijo panoge gradbeništvo in nepremičnine ter energetika in regulacija ogrevanja, prezračevanje in klimatizacije (v vsaki izmed skupin so po 3 respondenti navedli omenjene panoge). Na tretjem mestu pa sledijo podjetja, ki so dejavna na področju telekomunikacij, živilske industrije, gostinstva in turizma, trgovine na drobno in debelo, avto industrije ter IT in informatike (navedba po 2 respondentov v posamezni skupini). Sledili sta še dve panogi, kjer je bil le po en odgovor. Le po en odgovor sta prejeli komunalna panoga in panoga kemične industrije. Ostale panoge, kot so na primer zavarovalništvo, finance in bančništvo, mediji in založništvo, medicina in farmacija, izobraževanje ter pravna in odvetniška dejavnost, pa niso imele predstavnika v izvedenem spletnem anketiranju.

Tabela 6: Prikaz odgovorov respondentov glede panoge, v katero z vidika dejavnosti sodi podjetje, v katerem so zaposleni

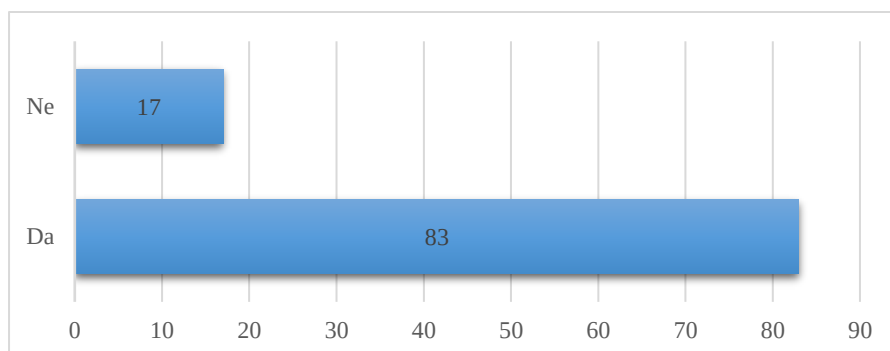
Izberite panogo v katero sodi vaše podjetje.				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Zavarovalništvo	0	0	0	0
Finance in bančništvo	0	0	0	0
Telekomunikacije	2	6	7	7
Logistika	5	14	17	23
Mediji in založništvo	0	0	0	23
Komunala	1	3	3	27
Živilska industrija	2	6	7	33
Medicina in farmacija	0	0	0	33
Izobraževanje	0	0	0	33
Gradbeništvo in nepremičnine	3	8	10	43
Pravna in odvetniška dejavnost	0	0	0	43
Gostinstvo in turizem	2	6	7	50
Trgovina na drobno in debelo	2	6	7	57
Energetika in regulacija ogrevanja, prezračevanje in klimatizacije	3	8	10	67
Avto industrija	2	6	7	74
Proizvodnja	5	14	17	91
IT in informatika	2	6	7	98
Kemična industrija	1	3	3	100
Skupaj	30	86	100	

Vir: lastno delo.

Vprašanja iz sklopa demografija so sicer predstavljala zadnja tri vprašanja spletne ankete, sicer sem na začetku anketiranja preverila pri vprašanju Q1, ali respondenti sploh poznajo

oziroma so že kdaj slišali za izraz »masovni podatki«. 6 respondentov (17 %) izraza ne pozna oziroma zanj še ni slišalo, 30 (83 %) pa jih je označilo, da izraz poznajo (slika 9).

Slika 9: Prikaz odgovorov respondentov glede poznavanje izraza »masovni podatki« (v %)



Vir: lastno delo.

To vprašanje nam glede na raziskovalno vprašanje naloge nudi možnost postavitve prve hipoteze, da statistično analiziramo dobljene odgovore (slika 10). Sicer ne gre za katero izmed dveh glavnih zastavljenih hipotez pričujoče naloge, pa vendar je smotrno preveriti ustreznost dobljenih podatkov in možnost sklepanja na populaciji.

H₀: Večina respondentov izraza »masovni podatki« ne pozna > 17.

H₁: Večina respondentov izraz »masovni podatki« pozna < ali = 17.

Slika 10: Preizkušanje hipoteze enake verjetnosti

Q1: Poznate ali ste že slišali za izraz »masovni podatki« (ang. Big Data)?

Odgovori	f ₀	f ₀ - Y _p	f _r
Da	30	29,5	18
Ne	6	6,5	18
Skupaj	36	36	

$$g=a-1=2-1=1 \quad \alpha=0,05$$

$$X^2 = \frac{(29,5-18)^2}{18} + \frac{(6,5-18)^2}{18} = 14,6944444$$

$$X^2 = 14,69444 > X^2 (\alpha=0,05; g= 1) = 3,8415$$

$$p - vrednost = 0,000126$$

Vir: lastno delo.

Izračunani X^2 je večji od teoretičnega, zato lahko H_0 zavrnilo in sprejmemo H_1 oziroma alternativno domnevo. S tveganjem, manjšim od $\alpha = 0,05$, trdimo, da izraz »masovni podatki« pozna večina oziroma več kot polovica respondentov ($p = 0,000126$). Podatki nakazujejo, da je zgolj 6 od 36 respondentov odgovorilo, da izraza ne poznajo, in tako je prispevek k X^2 največji na strani respondentov, ki so odgovorili z »Da«.

17 % respondentov, ki so odgovorili, da izraza »masovni podatki« ne poznajo, je nato prejelo vprašanje, kjer je bil izraz pojasnjen. Zanimalo me je ali menijo, da bi jih koncept masovnih podatkov v njihovem podjetju v prihodnosti morda zanimal. Na vprašanje so sicer odgovorili le 3 respondenti (od 6), kjer sta 2 (67 %) odgovorila, da jih masovni podatki tudi v prihodnosti ne bodo zanimali, in 1 (33 %), ki je odgovoril, da jih bodo zanimali (tabela 7). Nato so bili respondenti preusmerjeni na demografski sklop vprašanj, kjer so anketo tudi zaključili. Na vsa nadaljnja vprašanja, so tako odgovarjali le še respondenti, ki so pri prvem vprašanju odgovorili, da poznajo izraz »masovni podatki«.

Tabela 7: Prikaz odgovorov respondentov glede morebitne uporabe masovnih podatkov v prihodnosti v njihovem podjetju

Menite, da bi lahko predstavljeni koncept masovnih podatkov bil zanimiv za vaše podjetje v prihodnosti?				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Da	1	3	33	33
Ne	2	6	67	100
Skupaj	3	8	100	

Vir: lastno delo.

V nadaljevanju sem z vidika uporabe masovnih podatkov za notranjo rabo poslovanja podjetja zastavila nekaj vprašanj, ki se nanašajo dejansko na uporabo masovnih podatkov v velikih slovenskih podjetjih. Zanimali sta me je vključenost masovnih podatkov v strategijo podjetja ter merjenje in opažanje učinkov uporabe masovnih podatkov v poslovanju podjetja.

Na vprašanje o obdelavi masovnih podatkov v podjetju je odgovorilo 29 respondentov, kjer jih je večina, 21 respondentov, kar predstavlja 72 % vseh respondentov, odgovorila, da v njihovem podjetju obdelujejo masovne podatke (tabela 8).

Tabela 8: Obdelava masovnih podatkov v podjetju respondenta

Ali v vašem podjetju obdelujete t. i. masovne podatke?				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Da	21	58	72	72
Ne, vendar načrtujemo v prihodnosti	3	8	10	83
Ne	2	6	7	90
Ne vem	3	8	10	100
Skupaj	29	81	100	

Vir: lastno delo.

Med respondenti, ki poznajo izraz masovni podatki, ni pa nujno, da v njihovem podjetju že obdelujejo masovne podatke, me je zanimalo, ali masovni podatki oziroma aktivnosti v zvezi z masovnimi podatki predstavljajo del strategije njihovega podjetja. Od skupno 20 odgovorov je 13 respondentov, kar predstavlja večino s 65 %, odgovorilo, da imajo v njihovi strategiji podjetja vključene tudi masovne podatke (tabela 9).

Tabela 9: Odgovori respondentov glede vključenosti masovnih podatkov v strategijo njihovega podjetja

Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Da	13	36	65	65
Ne, vendar načrtujemo vključitev v prihodnosti	3	8	15	80
Ne	4	11	20	100
Skupaj	20	56	100	

Vir: lastno delo.

Respondenti, ki imajo, nimajo ali načrtujejo imeti v prihodnosti v strategijo podjetja vključene masovne podatke, so odgovarjali v nadaljevanju na vprašanje v zvezi s spremljanjem rezultatov izvajanja aktivnosti na področju masovnih podatkov. Pri izvajanju meritev spremljanja učinka se nakazujejo priložnosti za izboljšave, saj je 8 respondentov oziroma 40 % odgovorilo, da učinka ne merijo, prav tako 8 respondentov oziroma 40 % jih zaenkrat meritev še ne izvaja, vendar imajo to namen v prihodnosti in le 4 respondenti oziroma 20 % jih je odgovorilo, da meritve izvajajo (tabela 10).

Tabela 10: Odgovori respondentov glede merjenja učinka uporabe masovnih podatkov

Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Da	4	11	20	20
Ne, vendar nameravamo v prihodnosti	8	22	40	60
Ne	8	22	40	100
Skupaj	20	56	100	

Vir: lastno delo.

Respondentom, ki so na predhodno vprašanje glede izvajanja meritev učinka odgovorili pritrdilno, sem nato zastavila še vprašanje o rezultatu meritev. Od 4 respondentov, ki so odgovorili, da meritve izvajajo, so prav vsi odgovorili, da opažajo pozitiven učinek izvajanja aktivnosti na področju masovnih podatkov na poslovanje v njihovem podjetju (tabela 11).

Tabela 11: Rezultati meritev učinka uporabe masovnih podatkov

Opažate pozitiven ali negativen učinek na poslovanje?				
Kategorij	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Pozitiven	4	11	100	100
Negativen	0	0	0	100
Skupaj	4	11	100	

Vir: lastno delo.

V zasnovi oziroma v dispoziciji magistrskega dela sem zastavila namen, cilj, raziskovalno vprašanje ter hipoteze, ki so predstavljeni že v točkah 4.2 in 4.3, s katerimi želim preveriti dejansko stanje.

S prvo hipotezo sem želela preveriti, ali se večina velikih podjetij v Sloveniji poslužuje naprednih statističnih analiz masovnih podatkov, ki vključujejo tehnike napovedne analitike in podatkovnega rudarjenja. Tako sem v spletno anketo vključila vprašanje o dejanski uporabi omenjenih metod obdelave masovnih podatkov. Na vprašanje je odgovorilo skupaj 20 respondentov, kjer nisem uspela, kot je razvidno v nadaljevanju, statistično značilno dokazati (slika 9), da v večini velikih slovenskih podjetjih ne uporabljajo metod tekstovnega rudarjenja in napovedne analitike. Iz spodnjega prikaza pa je vseeno razvidno, da jih 8 oziroma 40 % metode tekstovnega rudarjenja še ni uporabilo, 5 oziroma 25 % pa jih ne ve. Podoben rezultat sem dobila tudi pri preverjanju glede uporabe metode napovedne analitike,

kjer je 9 respondentov oziroma 45 % odgovorilo, da metode niso uporabili, in 4 respondenti oziroma 20 %, da ne vedo (tabela 12).

Tabela 12: Odgovori glede uporabe tekstovnega rudarjenja in napovedne analitike v podjetjih anketiranih

Podatki omogočajo širok nabor različnih metod obdelave podatkov. Zanima nas ali ste že kdaj uporabili katero izmed spodaj navedenih metod.								
	Kategorije				Velja.	Št. enot	Povp.	Std. odklon
	Smo uporabili	Nismo uporabili	Ne vem	Skupaj				
Tekstovno rudarjenje (ang. Text mining)	7 (35 %)	8 (40 %)	5 (25 %)	20 (100 %)	20	36	1,9	0,8
Napovedna analitika (ang. Predictive analytics)	7 (35 %)	9 (45 %)	4 (20 %)	20 (100 %)	20	36	1,9	0,7

Vir: lastno delo.

Slika 11: Preizkušanje hipoteze enake verjetnosti za vprašanje Q5a: Tekstovno rudarjenje

Q5a: Tekstovno rudarjenje

Odgovori	f ₀	f ₀ - Y _p	f _r
Smo uporabili	7	6,5	6,7
Nismo uporabili	8	7,5	6,7
Ne vem	5	5,5	6,7
Skupaj	20		

$$g=a-1=3-1=2 \quad \alpha=0,05$$

$$X^2 = \frac{(6,5-6,7)^2}{6,7} + \frac{(7,5-6,7)^2}{6,7} + \frac{(5,5-6,7)^2}{6,7} = 0,3180$$

$$X^2 = 0,318 < X^2 (\alpha=0,05; g= 2) = 5,991$$

$$p - \text{vrednost} = 0,85535$$

Vir: lastno delo.

Izračunani X^2 je manjši od teoretičnega, zato ne moremo sprejeti alternativne domneve in trditi, da večina velikih slovenskih podjetij še ni uporabila metode tekstovnega rudarjenja. S tveganjem, manjšim od $\alpha = 0,05$, ne moremo trditi, da večina respondentov še ni uporabila tekstovnega rudarjenja ($p = 0,855345$). Podatki nakazujejo, da je 20 od 36 respondentov odgovorilo na vprašanje, kjer so bili odgovori razpršeni v tri kategorije, kjer ne moremo trditi, da je največji prispevek k X^2 na strani respondentov, ki so odgovorili z »Nismo uporabili« (slika 11).

Slika 12: Preizkušanje hipoteze enake verjetnosti za vprašanje Q5b: Napovedna analitika

Q5b: Napovedna analitika

Odgovori	f0	f0 - Yp	fr
Smo uporabili	7	6,5	6,7
Nismo uporabili	9	8,5	6,7
Ne vem	4	4,5	6,7
Skupaj	20		

$$g=a-1=3-1=2 \quad \alpha=0,05$$

$$X^2 = \frac{(6,5-6,7)^2}{6,7} + \frac{(8,5-6,7)^2}{6,7} + \frac{(4,5-6,7)^2}{6,7} = 1,2180$$

$$X^2 = 1,218 < X^2 (\alpha=0,05; g=2) = 5,991$$

$$p - \text{vrednost} = 0,54539$$

Vir: lastno delo.

Izračunani X^2 je manjši od teoretičnega, tako zavrnilo alternativno domnevo oziroma ne moremo trditi, da večina velikih slovenskih podjetij še ni uporabila metode napovedne analitike. S tveganjem, manjšim od $\alpha = 0,05$, ne moremo trditi, da večina respondentov še ni uporabila napovedne analitike ($p = 0,545392$). Podatki nakazujejo, da je 20 od 36 respondentov odgovorilo na vprašanje, kjer so bili odgovori razpršeni v tri kategorije, kjer ne moremo trditi, da je največji prispevek k X^2 na strani respondentov, ki so odgovorili z »Nismo uporabili« (slika 12).

Tako ne moremo sprejeti hipoteze, da večina velikih slovenskih še ni uporabila tekstovnega rudarjenja in napovedne analitike, oziroma zavrniti ničelne hipoteze, da je večina velikih slovenskih podjetij že uporabila tekstovno rudarjenje in napovedno analitiko.

Namen druge hipoteze je bil preveriti, ali si večina velikih podjetij v Sloveniji želi več analitičnih podatkov, ki bi jim pomagali izboljšati poslovanje z uporabo metod masovnih

podatkov. Tako sem v spletno anketo vključila vprašanje o področju vlaganja finančnih sredstev v različna področja uporabe masovnih podatkov. Med vnaprej podanimi odgovori sta bila dva, ki sta posredno povezana z večjimi količinami podatkov: zaledni sistemi (podatkovno skladišče), orodja za zajem in obdelavo podatkov. Za ta tip vprašanja, na osnovi katerega delam analizo, sem se odločila, ker sem želela hkrati preveriti dejansko pripravljenost in izvajanje naložb v proučevano področje, pri tem pa na osnovi odgovorov lahko sklepamo tudi, da podjetja, ki vlagajo v infrastrukturo, ki omogoča zajem in obdelavo večje količine podatkov, kar dejansko podana odgovora tudi predstavljata, razmišljajo in si želijo tudi večje količine podatkov, ki jih bodo lahko v prihodnosti obdelovali. Na vprašanje je skupno odgovorilo zgolj 10 respondentov oziroma 28 % vseh respondentov, kjer se za vnaprej podane odgovore: vizualizacija podatkov, zaposlitev novih sodelavcev z znanjem na področju napredne analitike in masovnih podatkov ter nič od navedenega ni odločil nihče, odgovor o vlaganjih v zaledne sisteme je izbralo 6 respondentov, kot zadnji pa je sledil odgovor, da so izvedli finančna vlaganja v orodja za zajem in obdelavo podatkov. Slednji odgovor so izbrali 4 respondenti (tabela 13).

Tabela 13: Prikaz odgovorov respondentov glede področja vlaganje finančnih sredstev v okviru masovnih podatkov

V kateri del vašega podjetja na področju masovnih podatkov ste vložili največ?				
Kategorije	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Vizualizacija podatkov	0	0	0	0
Zaledni sistemi (na primer: podatkovno skladišče)	6	17	60	60
Orodja za zajem in obdelavo podatkov	4	11	40	100
Zaposlitev novih sodelavcev z znanjem na področju napredne analitike in masovnih podatkov	0	0	0	100
Nič od navedenega	0	0	0	100
Skupaj	10	28	100	

Vir: lastno delo.

V nadaljevanju je izračunan X^2 , ki je manjši od teoretičnega, tako zavrnilo alternativno domnevo. S tveganjem, manjšim od $\alpha = 0,05$, ne moremo trdimo, da si večina respondentov želi večje količine podatkov ($p = 0,75062$). Podatki nakazujejo, da je 10 od 36 respondentov odgovorilo na vprašanje, kjer so bili odgovori razpršeni med dva odgovora, ter da je največji prispevek k X^2 na strani respondentov, ki so odgovorili z »Zaledni sistemi« (na primer: podatkovno skladišče) (slika 13).

Tako ne moremo sprejeti hipoteze, da si večina velikih slovenskih podjetij želi več analitičnih podatkov.

Slika 13: Preizkušanje hipoteze enake verjetnosti za Q10: V kateri del vašega podjetja na področju masovnih podatkov ste vložili največ?

Q10: V kateri del vašega podjetja na področju masovnih podatkov ste investirali največ?

Odgovori	f_0	$f_0 - Y_p$	f_T	$g=a-1=5-1=4$ $\alpha=0,05$
1 Vizualizacija podatkov	0			
2 Zaldeni sistemi (npr.: podatkovno skladišče)	6	5,5	5	
3 Orodja za zajem in obdelavo podatkov	4	4,5	5	
4 Zaposlitev novih sodelavcev na področju masovnih podatkov	0			
5 Nič od navedenega	0			
Skupaj	10			

$$X^2 = 0,5 < X^2(\alpha=0,05; g=4) = 9,49 \quad X^2 = \frac{(5,5-4)^2}{5} + \frac{(4,5-5)^2}{5} = 0,5$$

p - vrednost = 0,75062

Vir: lastno delo.

Namen naloge je proučiti obstoječe stanje, ozaveščenost in naklonjenost masovnim podatkom v velikih slovenskih podjetjih s ciljem preveriti, ali obstajajo neizkoriščene poslovne priložnosti. Spletna anketa v celoti sloni na zastavljenem namenu in cilju, pri čemer pa z vprašanjem o področju, na katero želijo podjetja vplivati z aktivnostmi masovnih podatkov, dobimo še dodatno pojasnjeno stanje. Od skupno 19 podanih odgovorov v tabeli 14 vidimo, da so respondenti, ki so sicer tukaj lahko označili več kot en odgovor, največ izbirali odgovor »Izboljšanje poslovnih procesov«, in sicer je bil odgovor izbran 13-krat. Sledi odgovor »Znižanje operativnih stroškov«, ki je bil izbran 10-krat, in nato še odgovor »Večja konkurenčnost podjetja«, ki je bil izbran 9-krat (tabela 14). Predvsem bi lahko zaključili, tudi če pogledamo izbor ostalih odgovorov, ki so bili sicer manjkrat izbrani, da podjetja zaenkrat masovne podatke uporabljajo predvsem za notranje poslovanje podjetja, ne pa tudi toliko za zunanje. Na primer vidimo, da je bil odgovor »Monetiziranje masovnih podatkov« izbran zgolj 1-krat in odgovor »Preoblikovanje in repozicioniranje podjetja« le 3-krat, kjer vidimo, da obstajajo priložnosti, da podjetja pričnejo tudi z uporabo masovnih podatkov za snovanje novih izdelkov in storitev na osnovi masovnih podatkov.

Tabela 14: Prikaz odgovorov respondentov glede področij, na katera želijo vpliva v okviru masovnih podatkov

Na katera področja vašega podjetja poskušate vplivati z aktivnostmi v okviru koncepta masovnih podatkov?					
Kategorije	Enote				
	Frekvence	Veljavni	Veljavni (v %)	Ustrezni	Ustrezni (v %)
Znižanje operativnih stroškov podjetja	10	19	53	36	28

se nadaljuje

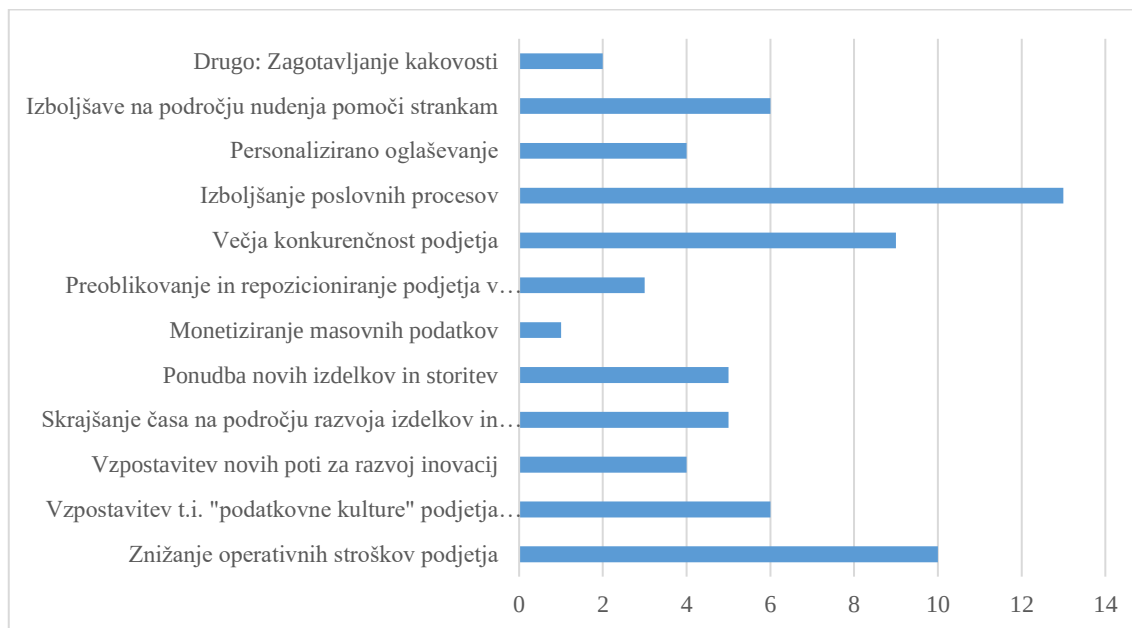
Tabela 14: Prikaz odgovorov respondentov glede področij, na katera želijo vpliva v okviru masovnih podatkov (nad.)

Na katera področja vašega podjetja poskušate vplivati z aktivnostmi v okviru koncepta masovnih podatkov?					
	Enote				
Kategorije	Frekvence	Veljavni	Veljavni (v %)	Ustrezni	Ustrezni (v %)
Vzpostavitev t. i. »podatkovne kulture« podjetja (ang. data-driven culture)	6	19	32	36	17
Vzpostavitev novih poti za razvoj inovacij	4	19	21	36	11
Skrajšanje časa na področju razvoja izdelkov in storitev	5	19	26	36	14
Ponudba novih izdelkov in storitev	5	19	26	36	14
Monetiziranje masovnih podatkov	1	19	5	36	3
Preoblikovanje in repozicioniranje podjetja v prihodnosti	3	19	16	36	8
Večja konkurenčnost podjetja	9	19	47	36	25
Izboljšanje poslovnih procesov	13	19	68	36	36
Personalizirano oglaševanje	4	19	21	36	11
Izboljšave na področju nujenja pomoči strankam	6	19	32	36	17
Drugo:	2	19	11	36	6
Skupaj		19		36	

Vir: lastno delo.

Še bolj nazoren in hiter pregled trenutnega stanja področij, na katera želijo vplivati podjetja, je prikazan na sliki 14, kjer dolžina vodoravnih stolpičev prikazuje z najdaljšim stolpičem največkrat izbrano področje in odgovor z najkrajšim stolpičem najmanjkrat izbran odgovor kot področje usmeritve podjetja z vidika masovnih podatkov.

Slika 14: Grafični prikaz odgovorov respondentov, na katera področja želijo vplivati z masovnimi podatki



Vir: lastno delo.

Slika 15: Preizkušnja hipoteze enake verjetnosti za vprašanje Q11: Na katera področja vašega podjetja poskušate vplivati z masovnimi podatki?

Q11: Na katera področja vašega podjetja poskušate vplivati z aktivnostmi v okviru koncepta masovnih podatkov?

Odgovori	f_0	F_t
Znižanje operativnih stroškov	10	5,58333
Vzpostavitev t.i. "podatkovne kulture" podjetja	6	5,58333
Vzpostavitev novih poti za razvoj inovacij	4	5,58333
Skrajšanje časa na področju razvoja izdelkov in storitev	5	5,58333
Ponudba novih izdelkov in storitev	5	5,58333
Monetiziranje masovnih podatkov	1	5,58333
Preoblikovanje in repozicioniranje podjetja v prihodnosti	3	5,58333
Večja konkurenčnost	9	5,58333
Izboljšanje poslovnih procesov	13	5,58333
Personalizirano oglaševanje	4	5,58333
Izboljšave na področju nudenja pomoči strankam	6	5,58333
Drugo: Zagotavljanje kakovosti	1	5,58333
Skupaj	67	

$$g = a - 1 = 12 - 1 = 11$$

$$\alpha = 0,05$$

$$X^2 = 25,18 > X^2(\alpha = 0,05; g = 11) = 19,68$$

$$p\text{-vrednost} = 0,00841$$

Vir: lastno delo.

Izračunani X^2 je večji od teoretičnega, zato lahko H_0 zavrnilo in sprejmemo H_1 oziroma alternativno domnevo. S tveganjem, manjšim od $\alpha = 0,05$, trdimo, da so dobljeni podatki reprezentativni oziroma da so zgolj 3 vnaprej podani odgovori od skupno 12 bili 32-krat izbrani, kar predstavlja slabo polovico vseh odgovorov, ki jih je bilo skupno 67 ($p =$

0,00841). Podatki nakazujejo, da je prispevek k X^2 največji na strani respondentov, ki so odgovorili z »Izboljšanje poslovnih procesov«, »Znižanje operativnih stroškov« in »Večja konkurenčnost« (slika 15).

Glede na pridobljene analitične rezultate raziskave kaže možnost sklepati, da velika slovenska podjetja uporabljajo masovne podatke za izboljšanje notranjega poslovanja podjetja. Zato me je tudi v nadaljevanju nekoliko bolj podrobno zanimalo, ali obstajajo tudi statistično značilne zveze med spremenljivkami. Tako v nadaljevanju z uporabo Pearsonovega korelacijskega koeficienta najprej preverjam povezavo med spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in »Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?« (tabela 15).

Tabela 15: Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in spremenljivko »Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?«

	Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?
Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	1	0,967
Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?	0,967	1

Vir: lastno delo.

Spremenljivki sta zelo močno pozitivno povezani, saj znaša Pearsonov korelacijski koeficient 0,967. Iz tega lahko sklepamo, da v podjetjih, ki imajo v strategiji vključene masovne podatke, tudi spremljajo in merijo učinek uporabe masovnih podatkov na rezultate podjetja (tabela 15).

Naprej, še nekoliko bolj podrobno nato preverim, kakšna je povezava med rezultatom merjenega učinka na poslovanje ter vključenostjo masovnih podatkov v strategijo (tabela 16).

Tabela 16: Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in spremenljivko »Opažate pozitiven ali negativen učinek na poslovanje?«

	Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	Opažate pozitiven ali negativen učinek na poslovanje?
Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	1	0,367
Opažate pozitiven ali negativen učinek na poslovanje?	0,367	1

Vir: lastno delo.

Povezanost med spremenljivkama »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in »Opažate pozitiven ali negativen učinek na poslovanje?« je pozitivna in šibka, saj znaša Pearsonov korelacijski koeficient 0,367. Iz dobljenega rezultata še vedno lahko sklepamo, da kjer predstavljajo masovnih podatki del strategije, ima uporaba masovnih podatkov pozitiven vpliv na poslovanje podjetja, vendar pa ni nujno, da je učinek na poslovanje pozitiven (tabela 16).

Kakšno vlogo ima povezanost podpore vodstva za vključenost masovnih podatkov v strategijo podjetja, sem preverila v nadaljevanju (tabela 17).

Tabela 17: Pearsonov korelacijski koeficient za spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in spremenljivko »Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?«

	Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?
Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?	1	0,90
Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?	0,90	1

Vir: lastno delo.

Povezanost med spremenljivkama »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in »Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?« je zelo močna, saj znaša Pearsonov

korelacijski koeficient 0,90. Iz dobljenega rezultata lahko sklepamo, da kjer predstavljajo masovni podatki del strategije, sta naklonjenost in podpora vodstva aktivnostim na področju masovnih podatkov zelo veliki oziroma si tudi vodstvo prizadeva za vključenost aktivnosti na področju masovnih podatkov v strategijo podjetja (tabela 17).

Kot navajajo Alharthi, Krotov in Bowman (2017, str. 288), je eden od glavnih izzivov organizacij, ki se ukvarjajo z masovnimi podatki, pomanjkanje znanja. Hkrati pa izhajam tudi iz že zapisanega, da iz dobljenih rezultatov analize ugotavljam, da se podjetja z vidika masovnih podatkov ukvarjajo pretežno z notranjim poslovanjem podjetja, kamor sodijo tudi organizacija, kultura podjetja in znanje zaposlenih. Zato sem na osnovi vprašanj o uporabi masovnih podatkov v podjetju in srečanju z izzivom pomanjkanja znanja izvedla izračun Pearsonovega korelacijskega koeficienta, da bi ugotovila, ali obstaja morda povezanost med uporabo masovnih podatkov v podjetjih in znanjem (tabela 18).

Tabela 18: Pearsonov korelacijski koeficient med spremenljivko »Ali v vašem podjetju obdelujete t. i. masovne podatke?« in spremenljivko »... pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem ...?«

	Ali v vašem podjetju obdelujete t. i. masovne podatke?	... pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem ...?
Ali v vašem podjetju obdelujete t. i. masovne podatke?	1	0,24
... pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem ...?	0,24	1

Vir: lastno delo.

Povezanost med spremenljivkama obstaja, vendar je šibka, saj znaša Pearsonov korelacijski koeficient 0,24. Tako lahko sklepamo, da je v določenih podjetjih vzrok, zakaj ne obdelujejo masovnih podatkov, tudi pomanjkanje znanja zaposlenih, vendar pa glede na rezultat statistične analize ne gre trditi, da je pomanjkanje znanja ključen dejavnik, ki bi vplival na odločitve o izvedbi aktivnosti na področju masovnih podatkov (tabela 18).

Zanimalo me je tudi, ali se v podjetjih srečujejo s težavo oziroma izzivom pomanjkanja znanja na področju masovnih podatkov, saj kot navajajo Alharthi, Krotov in Bowman (2017, str. 288), je to eden izmed glavnih izzivov podjetij. Od skupno 19 podanih odgovorov v tabeli 19 vidimo, da so respondenti največkrat izbirali odgovor »Da«, in sicer je bil odgovor izbran 16-krat, s 3 izbirami pa sledi odgovor »Ne« (tabela 19). Torej je kar 84 % vseh respondentov odgovorilo, da se soočajo z izzivom pomanjkanja znanja na področju masovnih podatkov.

Tabela 19: Prikaz odgovorov respondentov glede pomanjkanja znanja na področju masovnih podatkov

Alharthi, Krotov in Bowman ugotavljajo (2017, str. 288), da je eden od glavnih izzivov organizacij, ki se ukvarjajo z masovnimi podatki, pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem na področju analitike. Se tudi v vašem podjetju soočate s tem izzivom?				
Odgovori	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Da	16	44	84	84
Ne	3	8	16	100
Skupaj	19	53	100	

Vir: lastno delo.

Kot je razvidno iz tabele, lahko izpostavimo dva izziva, ki z vidika pogostosti izbora vodita med vnaprej podanimi odgovori. Na osnovi izbora respondentov tako vidimo, da jih je 9 oziroma 47 % izbralo odgovor »Zagotovitev kakovosti podatkov«, ki mu s 6 izbori oziroma 32 % sledi izziv »Zagotovitev upravljanja s podatki« (tabela 20).

Tabela 20: Prikaz odgovorov respondentov glede izzivov, s katerimi se soočajo

Med naštetimi trditvami, prosim, označite najbolj pomemben izziv za vas, s katerim se soočate v okviru masovnih podatkov.				
Odgovori	Frekvenca	Odstotek	Veljavni (v %)	Kumulativa (v %)
Zagotovitev kakovostnih podatkov	9	25	47	47
Varovanje zasebnosti podatkov	2	6	11	58
Zagotovitev varnosti podatkov	1	3	5	63
Zagotovitev finančnih sredstev	1	3	5	68
Zagotovitev upravljanja s podatki	6	17	32	100
Drugo:	0	0	0	100
Skupaj	19	53	100	

Vir: lastno delo.

5 DISKUSIJA

V poglavju diskusije bom rezultate raziskave primerjala s teoretičnimi izhodišči in že izvedenimi vprašalniki, ki sem jih uporabila za sestavo spletne ankete, ter tudi predstavljenimi spoznanji iz teoretičnega dela, ki jih navajajo znanstveni in strokovni avtorji.

Vseh zastavljenih 19 vprašanj bom združila v smiselne sklope, na osnovi katerih bom opredelila trditev o spoznanjih in ugotovitvah.

5.1 V Sloveniji analizirajo masovne podatke velika podjetja, ki izhajajo iz storitvene in proizvodne dejavnosti

V raziskavi pričujoče naloge je bila večina respondentov predstavnikov logistične (17 %) in proizvodne (17 %) dejavnosti, ki jima sledijo gradbeništvo in nepremičnine (10 %), energetika in regulacija ogrevanja, prezračevanje in klimatizacije (10 %).

Iz raziskave, ki so jo izvedli v obdobju od leta 2011 do leta 2016 na globalni ravni, izhaja, da so najbolj aktivna podjetja na področju masovnih podatkov podjetja, ki nudijo storitveno dejavnost in uporabljajo lokacijske podatke, ter tudi podjetja v maloprodaji (Henke in drugi, 2016, str. 29–30). Leta 2015, ko je bila izvedena raziskava o uporabi masovnih podatkov v Sloveniji, so na SURS ugotovili, da je bilo z vidika dejavnosti 14 % takih, ki se ukvarjajo s storitveno dejavnostjo, in 8 % takih, ki se ukvarjajo s proizvodno dejavnostjo (Zupan, 2016b).

Na vprašanja v sklopu demografskih vprašanj, kamor uvrščam tudi vprašanje o dejavnosti podjetja, iz katerega so prihajali respondenti, so odgovarjali vsi respondenti, ne glede na to, ali so masovne podatke poznali ali ne. Tako ni bilo možno brez dodatne obdelave podatkov zaključiti, da je v primeru izvedene spletne ankete največ podjetij, ki se ukvarjajo z analizo masovnih podatkov, iz storitvene ali proizvodne dejavnosti, tako kot so to zaključili Henke in drugi (2016) ter tudi Zupan (2016a) v svojih raziskavah. Po obdelavi podatkov in podrobnejši analizi, kjer sem najprej primerjala izbor odgovorov pri 1. vprašanju o poznavanju masovnih podatkov (Da = 1 in Ne = 2), naprej pri tistih, ki so izbrali »Da«, sem pri 3. vprašanju primerjala izbor odgovorov (Da = 1) z izborom odgovorov pri 16. vprašanju, kjer so respondenti izbrali »Logistika« (Logistika = 4). Enak postopek primerjave sem izvedla tudi za odgovor »proizvodnja«, kjer pa sem le za razliko od prejšnjega odgovora primerjala 1. in 3. vprašanje oziroma odgovore z navedbami, ki so jih respondenti sami zapisali pod »Drugo«.

Na osnovi svojih ugotovitev lahko trdim, da so velika podjetja v Sloveniji konec leta 2017 z vidika dejavnosti, s katero se ukvarjajo, najbolj aktivna na področju analize masovnih podatkov, tako na področju storitvene kot tudi proizvodnje dejavnosti, za razliko od ugotovitev iz sekundarnih virov, kjer sem ugotovila, da na globalni ravni izkoriščajo možnosti masovnih podatkov v večji meri v podjetjih s storitveno dejavnostjo, za razliko od ugotovitev Zupana (2016b), ki je raziskavo izvajal v letu 2015 in prav tako ugotovil, da je takrat bila večina podjetij s področja storitvene dejavnosti.

5.2 Masovne podatke pozna večina velikih podjetij v Sloveniji, veliko jih te podatke tudi obdeluje

Po ugotovitvah v raziskavi pozna masovne podatke med respondenti velikih podjetij v Sloveniji 83 % vprašanih, od tega jih 58 %, kar predstavlja večino, masovne podatke tudi obdeluje.

V raziskavi, ki jo je izvedel Zupan (2016b) leta 2015, je bilo ugotovljeno, da masovne podatke analizira 11 % vprašanih, kjer pa niso bila vključena v vzorec zgolj velika slovenska podjetja, temveč podjetja, ki imajo 10 in več zaposlenih. Od skupno 11 % podjetij, ki so analizirala masovne podatke že takrat, pa je bilo 38 % velikih podjetij, kar je bila večina.

Zaključim lahko, da se analize masovnih podatkov v največji meri lotevajo v velikih podjetjih v Sloveniji. Trditev lahko postavim na osnovi spoznanj iz sekundarnih virov, kot tudi iz izvedene empirične raziskave.

5.3 Ne morem trditi, da se velika podjetja v Sloveniji ne poslužujejo tehnik tekstovnega rudarjenja in napovedne analitike

V nalogi sem postavila dve hipotezi, ki sem jih poskušala preveriti in dokazati. Prva hipoteza je, ali se večina velikih podjetij v Sloveniji poslužuje naprednih statističnih analiz masovnih podatkov, ki vključujejo tehnike napovedne analitike in podatkovnega rudarjenja.

Na vprašanje je odgovorilo zgolj 20 respondentov, kjer so sicer ob pregledu opisnih statistik predstavljali večino odgovori, ki so bili, da proučevanih metod še niso uporabili. Z izvedbo statistične analize pa vseeno nisem uspela dokazati, da je rezultat statistično značilen in da v večini velikih slovenskih podjetij ne uporabljajo metod tekstovnega rudarjenja in napovedne analitike.

Kot je razvidno iz tabele o primerjavi različnih tehnik analize masovnih podatkov (tabela 2), ki so jo pripravili Yaqoob in drugi (2016, str. 1239), ter če izhajam iz spoznanja, da je večina naših respondentov, ki izvajajo analize masovnih podatkov, predstavnikov storitvene in proizvodne dejavnosti, bi se izbira kakšnih drugih tehnik morda odražala v drugačnem izidu odgovorov.

Tako ne morem trditi, da se velika podjetja v Sloveniji ne poslužujejo različnih tehnik analize masovnih podatkov, prav tako statistično značilno ne morem trditi, da se ne poslužujejo tekstovnega rudarjenja in napovedne analitike.

5.4 Uporaba masovnih podatkov predstavlja enega izmed poslovnih procesov, za katere skrbijo v velikih podjetjih v Sloveniji

– Vključenost v strategijo podjetja in spremljanje učinkov

Ugotovila sem, da ima 13 od skupno 20 respondentov, kar predstavlja večino s 65 %, v strategiji podjetja vključene tudi masovne podatke.

Spremljanje učinka izvajanja analiz masovnih podatkov je nekoliko slabše oziroma učinek spremlja zgolj 8 respondentov oziroma 40 %, prav tako 8 respondentov oziroma 40 % zaenkrat meritev še ne izvaja, vendar imajo to namen početi v prihodnosti. Le 4 respondenti oziroma 20 % jih je odgovorilo, da meritve izvajajo. Iz tega izhaja, da obstajajo priložnosti za podjetja, da v aktivnosti vključijo tudi meritve učinkov izvajanja notranjih procesov na poslovanje podjetja, konkretno tukaj analiziranje masovnih podatkov.

Zanimalo me je, ali obstajajo statistično značilne korelacije med spremenljivkami. Tako sem z uporabo Pearsonovega korelacijskega koeficienta preverila povezavo med spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in »Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?«. Ugotovila sem, da sta spremenljivki zelo močno pozitivno povezani. To pomeni, da lahko trdim, da v podjetjih, ki imajo v strategijo vključene masovne podatke, tudi spremljajo in merijo učinek uporabe masovnih podatkov na rezultate podjetja.

Podjetja se vse bolj odločajo za vodenje podjetja in sprejemanje odločitev na osnovi zbranih, obdelanih analiziranih podatkov z namenom ustvarjanja bolj usmerjenih oziroma specializiranih storitev, zniževanja tveganj in izboljšanja učinkovitosti poslovanja podjetja (Fan, Han & Liu, 2014, str. 296). Po navedbi LaValle, Lesserja, Shockley-ja, Hopkinsa in Kruschwitza (2011, str. 23) najbolj uspešna podjetja uporabljajo analitiko petkrat pogosteje kot podjetja, ki so manj uspešna. Za ugotovitev o uspešnosti in vplivu uporabe in izvajanja analiz masovnih podatkov na uspešnost poslovanja podjetja je treba vzpostaviti tudi mehanizem spremljanja učinka ali vzpostavitev izvajanja meritev učinkovitosti poslovnega procesa.

Na osnovi lastne raziskave trdim, da v velikih podjetjih v Sloveniji, kjer imajo vključene masovne podatke v strategijo podjetja, v večini podjetij njihovo upravljanje in izvajanje tudi spremljajo in merijo. Tega pa ne morem trditi tudi za tiste, ki masovnih podatkov nimajo vključenih v strategijo. Ugotovila sem, da je na splošno merjenje učinka na poslovanje precej nizko. Tako obstaja priložnost za izboljšave na tem področju.

– Finančni vložek

Z drugo hipotezo sem želela preveriti, ali večina velikih podjetij v Sloveniji želi več analitičnih podatkov, ki bi jim pomagali izboljšati poslovanje z uporabo metod masovnih podatkov.

Kot odgovor na zastavljeno hipotezo sem uporabila vprašanje, ki ni neposredno zasnovano oziroma vezano na področje, temveč sem želela prek vprašanja o finančnem vlaganju ugotoviti, ali si želijo več podatkov. Izhajala sem iz domneve, da če podjetje želi več podatkov, to pomeni, da se mora tudi ustrezno nanje pripraviti oziroma poskrbeti z ustrezno infrastrukturo v smislu podatkovnega skladišča, zadostnih kapacitet hranjenja podatkov, kjer za slednje potrebuje zmogljivejše strežnike oziroma večje kapacitete. Tako sem pri vprašanju o področjih njihovega vlaganja ugotovila, da je skupno odgovorilo zgolj 10 respondentov oziroma 28 % vseh respondentov. Odgovor o vlaganjih v zaledne sisteme je izbralo največ respondentov (6), sledil je odgovor, da so izvedli finančna vlaganja v orodja za zajem in obdelavo podatkov. Slednji odgovor so sicer izbrali zgolj 4 respondenti. Ostalih vnaprej ponujenih odgovorov ni izbral noben respondent.

Henke in drugi (2016, str. 29-30) ugotavljajo, da je možno v okvirno enem desetletju podvojiti prihodke in prihranke v poslovanju podjetja od naložb v podatkovne in analitične zmogljivosti. V zadnjih 5 letih so se podjetja lotila pomembnih vlaganj v okviru masovnih podatkov (Davenport & Bean, 2017, str. 7).

Na osnovi zbranih podatkov ne morem trditi, da si večina velikih slovenskih podjetij želi več analitičnih podatkov. Vseeno pa glede na zbrane podatke (določena podjetja so vložila finančna sredstva na področju masovnih podatkov) zaznavam, da nekaj podjetij pričakuje dodatne možnosti, s katerimi bodo lahko ustvarila pozitivne učinke na poslovanje podjetja.

– Podjetja želijo primarno izboljšati notranje poslovanje in tudi konkurenčno prednost

Od skupno 19 podanih odgovorov pri vprašanju o področju, na katero želijo podjetja vplivati z masovnimi podatki, so respondenti največkrat izbirali odgovor »Izboljšanje poslovnih procesov«, in sicer je bil odgovor izbran 13-krat, sledi odgovor »Znižanje operativnih stroškov«, ki je bil izbran 10-krat, in nato še odgovor »Večja konkurenčnost podjetja«, ki je bil izbran 9-krat.

Lee (2017, str. 299) meni, da je na področju masovnih podatkov možno vplivati na različna področja v podjetju. Davenport in Bean (2017, str. 8) navajata primere področij, na katera je možno vplivati. Navedena področja so zniževanje operativnih stroškov, vzpostavitev novih poti, monetiziranje masovnih podatkov, večja konkurenčnost ipd. Kubina, Varmus in Kubinova (2015, str. 564) menijo tudi, da podjetja, ki ne izkoriščajo prednosti masovnih

podatkov oziroma se ne poslužujejo uporabe masovnih podatkov, s tem dajejo na trgu prednost svojim konkurentom.

Iz dobljenih rezultatov, glede na teoretično osnovo, bi lahko trdila, da velika podjetja v Sloveniji želijo z masovnimi podatki vplivati predvsem na izboljšanje oziroma učinkovitost notranjega poslovanja podjetja ter da želijo tudi izboljšati svojo konkurenčnost na trgu.

– **Naklonjenost vodstva**

Z vidika naklonjenosti vodstva upravljanju masovnih podatkov v podjetju me je zanimalo predvsem, ali obstaja povezanost med spremenljivko »Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?« in »Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?«. Iz analize sem ugotovila, da je povezanost med spremenljivkama zelo visoka; Pearsonov korelacijski koeficient znaša 0,90.

Podjetja v obdobju masovnih podatkov ne uspejo zgolj, ker imajo več in boljše podatke, temveč ker imajo vodstvene time, ki postavljajo jasna pričakovanja, merljive cilje, jasno opredeljujejo, kaj pomeni uspeh, in zastavljajo prava vprašanja (McAfee & Brynjolfsson, 2012).

Iz dobljenega rezultata lahko sklepamo, da sta, kjer predstavljajo masovni podatki del strategije, naklonjenost in podpora vodstva aktivnostim na področju masovnih podatkov zelo velika oziroma si tudi vodstvo prizadeva za vključenost aktivnosti na področju masovnih podatkov v strategijo podjetja.

– **Znanje v podjetju in drugi izzivi**

Podjetjem primanjkuje kadra z ustreznim znanjem na področju analitike masovnih podatkov, kot ugotavljam v analizi pridobljenih podatkov v spletni anketi. Kar 84 % od vseh respondentov je odgovorilo, da se soočajo z izzivom pomanjkanja znanja.

Zanimala me je tudi povezanost med uporabo masovnih podatkov v podjetjih in znanjem, torej, ali se z izzivom pomanjkanja znanja srečujejo v podjetjih, kjer uporabljajo masovne podatke. Zato sem izračunala Pearsonov korelacijski koeficienta, da bi ugotovila, ali obstaja povezanost med omenjenima spremenljivkama. Ugotovila sem, da povezanost sicer obstaja, vendar je precej nizka.

Kot ugotavljajo Alharthi, Krotov in Bowman (2017, str. 288), je eden od glavnih izzivov organizacij, ki se ukvarjajo z masovnimi podatki, pomanjkanje znanja na področju analitike.

Na osnovi podatkov iz ankete ugotavljam tudi, da se v podjetjih poleg izziva »pomanjkanje znanja« soočajo tudi z izzivom zagotavljanja kakovostnih podatkov in upravljanja podatkov.

Ugotavljam, da, izhajajoč iz sekundarnih virov, predstavlja pomanjkanje znanja na področju upravljanja masovnih podatkov enega od večjih izzivov. To sem tudi ugotovila na osnovi izvedene ankete, kjer je večina respondentov navedla, da se soočajo z izzivom pomanjkanja znanja. Ne morem pa trditi, da je pomanjkanje znanja ključen dejavnik, ki bi vplival na odločitev o obdelavi oziroma upravljanju masovnih podatkov, saj je bila korelacija med spremenljivkama zelo nizka.

SKLEP

Namen magistrskega dela je bil proučiti ozaveščenost, obstoječe stanje in naklonjenost velikih podjetij v Sloveniji uporabi masovnih podatkov. Kot cilj magistrskega dela sem si zastavila preveriti, ali v velikih slovenskih podjetjih obstajajo neizkoriščene poslovne priložnosti na področju masovnih podatkov.

Na osnovi pregleda sekundarnih virov, povzetih znanstvenikov in strokovnjakov na področju masovnih podatkov na globalni ravni ter analize podatkov, pridobljenih z izvedbo spletne ankete, sem v sklopu diskusije predstavila ugotovitve. S spletno anketo sem pridobila vpogled v stanje v velikih slovenskih podjetjih na področju uporabe masovnih podatkov. K sodelovanju v spletni anketi sem prek elektronskega sporočila povabila vsa velika slovenska podjetja, od katerih sem uspela pridobiti elektronski naslov. Na osnovi analiziranih podatkov, dobljenih s spletnim anketiranjem, sem v diskusiji, prvotno 19 zastavljenih vprašanj, združila v 4 smiselne sklope, na osnovi katerih lahko predstavimo spoznanja in stanje v podjetjih.

V prvem sklopu so zbrana in analizirana vprašanja, vključno s teoretičnimi izhodišči s področja dejavnosti, s katero se podjetje ukvarja. Ugotovitve prvega sklopa so, da so velika podjetja v Sloveniji z vidika dejavnosti, s katero se ukvarjajo, najbolj aktivna na področju analize masovnih podatkov na dveh področjih, storitvenem in proizvodnem, kar je nekoliko v nasprotju z ugotovitvami iz sekundarnih virov, kjer podatki kažejo na največjo mero aktivnosti v storitveni dejavnosti. Kot sem sicer že zapisala v sklopu metodologije, pa naj tudi v sklepnem delu dodam, da pri zbiranju podatkov o dejavnosti podjetij, nisem izhajala iz standardne klasifikacije dejavnosti. Tako dopuščam možnost, da bi rezultat v primeru uporabe standardne klasifikacije dejavnosti morda bil drugačen.

V okviru drugega sklopa sem ugotovila, da v velikih podjetjih v Sloveniji v večini poznajo ter tudi že izvajajo analize masovnih podatkov. Ugotovitev izhaja tako iz sekundarno zbranih podatkov kot tudi iz primarne raziskave, kjer so respondenti v spletni anketi podali odgovor, da se v njihovem podjetju ukvarjajo z analizo masovnih podatkov. Kljub rezultatu pa je pomembno opozoriti, da se je spletne ankete udeležilo 36 respondentov od skupno vabljenih 238. To pomeni, da na osnovi primarno pridobljenih podatkov zagotovo težko sklepam zaključke in se zato v okviru sklepanja zaključkov dodatno naslanjam tudi na sekundarno zbrane in analizirane podatke.

V okviru tretjega sklopa na podlagi primarno zbranih podatkov, ne morem trditi, da se velika podjetja v Sloveniji ne poslužujejo različnih tehnik analize masovnih podatkov, prav tako ne morem trditi, da se ne poslužujejo tekstovnega rudarjenja in napovedne analitike. Zbrani podatki in rezultati analize, žal niso bili statistično značilni, kljub dejstvu, da je bilo z vidika števila podanih odgovorov največ takšnih, ki so odgovorili, da proučevanih metod še niso uporabili.

Zadnji, četrti sklop združuje rezultate več različnih področij z vidika poslovanja podjetja. Področja sem razdelila na: strateški in finančni vidik, vpliv na poslovne procese, naklonjenost vodstva ter znanje in drugi izzivi. Na osnovi analiziranih podatkov dobljenih z izvedbo spletne ankete, sem ugotovila, da v večini primerov velika podjetja, kjer upravljajo masovne podatke, te tudi spremljajo ter merijo njihov učinek. Vseeno pa glede na nizko udeležbo v spletnem anketiranju, je težko sklepati, da ugotovitev velja sklepati tudi na populaciji. Z posrednim vprašanjem, vezanim na finančni vidik, glede na višino finančnega vložka v preteklih 5 letih, sem preverjala željo podjetji po večji količini podatkov. Vzorec je bil nizek oziroma prenizek za statistično značilno sklepanje, vendar pa dopušča možnost sklepanja, da si večina velikih slovenskih podjetij želi več analitičnih podatkov. Tako ostaja na tem področju priložnost, za dodatno, bolj poglobljeno raziskavo z uporabo neposrednih vprašanj, ki nedvoumno omogočajo sklepanje zaključkov. Samokritično tudi ocenjujem, da bi bilo bolj smotno to vprašanje zastaviti bolj neposredno ter tako povečati možnost za boljši rezultat. Na osnovi preverjanja namena, ki ga zasledujejo z upravljanjem masovnih podatkov, sem ugotovila, da podjetja želijo vplivati na izboljšanje in učinkovitost notranjega poslovanja ter tudi na konkurenčnost. Podatki kažejo tudi na visoko stopnjo naklonjenosti vodstva masovnim podatkom, vendar pa sem tako na osnovi sekundarnih kot tudi primarnih podatkov ugotovila, da se podjetja soočajo s pomanjkanjem znanja s področja masovnih podatkov. Torej predstavlja pomanjkanje znanja poleg izziva zagotovitve kakovostnih podatkov in upravljanja podatkov tudi enega od ključnih izzivov, s katerimi se velika podjetja v Sloveniji soočajo.

Izhajajoč iz zastavljenega cilja, sem zasnovala raziskovalno vprašanje, kako dobro v velikih slovenskih podjetjih izkoriščajo možnosti za izboljšanje, prilagajanje ali širitev poslovanja podjetja na osnovi uporabe masovnih podatkov. Na osnovi ugotovitev, predvsem z vidika zmanjšanja izzivov in širitve namena uporabe masovnih podatkov, obstajajo priložnosti, da podjetja vplivajo na omenjena področja. Na osnovi prikazov v teoretičnem delu naloge, v sklopu sekundarnih podatkov, kot tudi v empiričnem delu, so prikazane priložnosti za uporabo masovnih podatkov v različne namene ter tudi pozitivni učinki njihove uporabe na različnih področjih. Ugotovila sem, da tudi z vidika znanja kadrov obstaja priložnost za izobraževalne institucije, saj so rezultati pokazali, da se podjetja soočajo s pomanjkanjem znanja. Izhajajoč iz ugotovitve menim, da bi lahko različne izobraževalne institucije ponudile strokovne in znanstvene tečaje ter certificiranja s področja upravljanja masovnih podatkov, kot so na primer podatkovni inženir, podatkovni znanstvenik, analitik masovnih podatkov ipd. Po pregledu sekundarnih virov sem namreč tudi zasledila, da v tujini že obstajajo možnosti za dodatno izobraževanje, naj si bo prek spleta ali v živo.

Z vidika možnosti uporabe masovnih podatkov za različne namene pa menim, da bi podjetja lahko te bolj uporabljala za napovedovanje poslovnih dogodkov, za še večjo konkurenčnost predvsem z vidika marketinških aktivnosti, osebno prilagojene ponudbe in izboljšanja zadovoljstva strank. Predlagala bi nadaljnje raziskovanje, predvsem z vidika marketinških aktivnosti, ki se zdijo z vidika priložnosti ustvarjanja dodatnih prihodkov ter hkrati prihrankov na osnovi zbranih in analiziranih podatkov, precej zanimiva za nadaljnji razvoj in izkoriščenje tudi na domačem, slovenskem trgu. Če izhajam še iz empiričnega dela naloge, kjer sem preverjala obstoječe stanje, glede uporabe napovedne analitike v podjetjih in kjer moram opozoriti, da rezultat sicer ni bil statistično značilen, lahko tudi tukaj opazimo, da večina sodelujočih v raziskavi, ki so odgovorili na vprašanje, tehnik napovedne analitike še niso uporabili.

Tako bi lahko odgovor na raziskovalno vprašanje zaključila, da so velika podjetja v Sloveniji že aktivna na področju uporabe masovnih podatkov ter da tudi že izkoriščajo priložnosti za izboljšanje notranjih procesov, vendar pa še obstaja priložnosti tako za prilagajanje poslovanja na področju notranjih in zunanjih procesov poslovanja kot tudi konkurenčnosti in širitvi poslovanja.

V prihodnosti je pričakovati dodatno rast podatkov zaradi vse hitrejšega razvoja t. i. interneta stvari, zato je tudi pričakovati, da bo povpraševanje po znanju in uporabi masovnih podatkov vse večje in vse bolj pomembno. Menim, da bodo aktivnosti podjetij na področju masovnih podatkov v prihodnosti še bolj pomembne, tako na strokovnem kot tudi znanstvenem področju.

LITERATURA IN VIRI

1. Akter, S., Wamba, S. A., Gunasekaran, A., Dubey, R. & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182(2016), 113–131.
2. Alharthi, A., Krotov, V. & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons*, 60(3), 285–292.
3. Bagnoli, V. (2015). Competition fort he Effectiveness of Big Data Benefits. *International Review of Intellectual Property and Competition Law*, 46(6), 629–631.
4. Braganza, A., Brooks, L., Nepelski, D., Ali, M. & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, 70(2017), 328–337.
5. Braunsberger, F., Hlavaty, M., Schlamberger, N. & Stevanovič, S. (2010). *Standardna klasifikacija dejavnosti 2008*. Pridobljeno 3. oktober 2018 iz <https://www.stat.si/doc/pub/skd.pdf>
6. Davenport, T. H., Barth, P. & Bean, R. (2012). How `Big Data` is Different. *MIT Sloan Management Review*, 54(1), 21–24.

7. Davenport, T. H. & Bean, R. (2017). *Big Data Executive Survey 2017. Big Data Business Impact: Achieving Business Results through Innovation and Disruption*. Boston: NewVantage Partners LLC.
8. Davenport, T. H. & Dyché, J. (2013). *Big Data in Big Companies*. Oregon: International Institute for Analytics.
9. de Waal-Montgomery, M. (2015, 15. januar). *World's data volume to grow 40 % per year & 50 times by 2020: Aureus*. Pridobljeno 27. aprila 2018 iz <https://e27.co/worlds-data-volume-to-grow-40-per-year-50-times-by-2020-aureus-20150115-2/>
10. Dublin City Council. (brez datuma). *About Dublin City Council*. Pridobljeno 26. novembra 2017 iz <http://www.dublincity.ie/main-menu-your-council/about-dublin-city-council>
11. Elgendy, N. & Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, 100(2016), 1071–1084.
12. Etihad Airways. (brez datuma). *Corporate profile*. Pridobljeno 26. novembra 2017 iz <http://www.etihad.com/en-si/about-us/corporate-profile/>
13. Fan, J., Han, F. & Liu, H. (2014). Challenges of Big Data analysis. *National Science Review*, 1(2014), 293–314.
14. Feng, K. & Chen, L. (2017). The Evolution of Artificial Ecosystem Based on the Ecological Chain Algorithm. *Advances in Modelling and Analysis B*, 60(1), 110-122.
15. Finlay, S. (2014). *Predictive Analytics, Data Mining and Big Data*. UK: Palgrave Macmillan.
16. Gepp, A., Linnenluecke, M. K., O'Neill, T. J. & Smith, T. (2018). Big data techniques in auditing research and practice: Current trends and future opportunities. *Journal of Accounting Literature*, 40(2018), 102–115.
17. Google Trends. (2018). *Iskanje po ključnih besedah: Big Data, IoT, Cloud, Cybersecurity in Augmented reality*. Pridobljeno 5. maja 2018 iz <https://trends.google.com/trends/explore?cat=12&date=2010-01-01%202018-01-01&q=Big%20Data,IoT,Cloud,Cybersecurity,Augmented%20reality>
18. Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M. & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of strategic information systems*, 26(2017), 191–209.
19. Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B. & Sethupathy, G. (2016). *The age of analytics: competing in a data-driven world*. New York: McKinsey Global Institute.
20. Intel IT Center (2012). *Peer Research: Big Data Analytics*. Santa Clara: Intel Corporation.
21. Kern Pipan, K., Bertok, J. & Kotnik, I. (2017). *Masovni podatki – velika priložnost za javno upravo – izkušnje pilotnega projekta*. Ljubljana: Ministrstvo RS za javno upravo.
22. Krotov, V. (2017). The Internet of Things and new business opportunities. *Business Horizons*, 60(6), 831–841.
23. Kubina, M., Varmus, M. & Kubinova, I. (2015). Use of big data for competitive advantage of company. *Procedia Economics and Finance*, 26(2015), 561–565.

24. LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, 52(2), 20–32.
25. Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(2017), 293–303.
26. Lee, I. & Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(2015), 431–440.
27. Malek, Y. N., Kharbouch, A., El Khoukhi, H., Bakhouya, M., De Florio, V., El Ouadghiri, D., Latre, S. & Blondia, C. (2017). On the use of IoT and Big Data Technologies for Real-time Monitoring and Data Processing. *Procedia Computer Science* 113(2017), 429–434.
28. Mantelero, A. (2017). Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework. *Computer Law & Security Review*, 33(2017), 584-602.
29. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, H. A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. New York: McKinsey Global Institute.
30. Mazzei, M. J. & Noble, D. (2017).: A framework for corporate strategy. *Business Horizons* 60(2017), 406-407.
31. McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, 90(10), 61-68.
32. Metallo, C., Agrifoglio, R., Schiavone, F. & Mueller, J. (2018). Understanding business model in the Internet of Things industry. *Technological Forecasting & Social Change*, (2018), 1–9.
33. Miloslavskaya, N. & Tolstoy, A. (2016). Big Data, Fast Data and Data Lake Concepts. *Procedia Computer Science*, 88(2016), 300–305.
34. Misra, A., Sharma, A., Gulia, P. & Bana, A. (2014). Big Data: Challenges and Opportunities. *International Journal of Innovative Technology and Exploring Engineering*, 4(2), 41.
35. O'Reilly, T. (2007). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Communication & strategies*, 65(1st quarter 2007), 17–37.
36. Özköse, H., Ari, E. S. & Gencer, C. (2015). Yesterday, Today and Tomorrow of Big Data. *Procedia – Social and Behavioral Sciences*, 195(2015), 1042–1050.
37. Pohorec, S. (2015a). Analizirajmo veliko podatkov! *Monitorpro*, 6, 38–41.
38. Pohorec, S. (2015b). Kako do uspešne strategije big data. *Monitorpro*, 6, 41–43.
39. Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187–195.
40. Saggi, M. K. & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing and Management*, 54(5), 758-790.

41. Santos, M. Y., e Sá, J. O., Andrade, C., Lima, F. V., Costa, E., Costa, C., Martinho, B. & Galvao, J. (2017). A Big Data system supporting Bosch Braga Industry 4.0 strategy. *International Journal of Information Management*, 37(6), 750–760.
42. Sivarajah, U., Kamal, M. M., Irani, Z. & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70(2017), 263–286.
43. Svet Evrope. (1994). *Konvencija o varstvu posameznika glede na avtomatsko obdelavo osebnih podatkov*. Pridobljeno 4. maja 2018 iz http://www.svetevrope.si/sl/dokumenti_in_publikacije/konvencije/108/
44. TechTarget IoT Agenda. (oktober, 2016). *Definition drone (unmanned aerial vehicle, UAV)*. Pridobljeno 3. maja 2018 iz <https://internetofthingsagenda.techtarget.com/definition/drone>
45. The Daimler Group. (brez datuma). *Company*. Pridobljeno 26. novembra 2017 iz <https://www.daimler.com/company/>
46. The Walt Disney Company. (brez datuma). *About*. Pridobljeno 26. novembra 2017 iz <https://thewaltdisneycompany.com/about/>
47. United Parcel Services. (brez datuma). *About UPS*. Pridobljeno 26. novembra 2017 iz <https://www.ups.com/at/en/about.page?>
48. Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B. & Vasilakos, A. V. (2016). Big data: From beginning to future. *International Journal of Information Management*, 36(6), 1231–1247.
49. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R. & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70(2017), 356–365.
50. Qin, Y., Sheng, Q. Z., Falkner, N. J. G., Dustdar, S., Wang, H. & Vasilakos, A. V. (2016). When things matter: A survey on data-centric internet of things. *Journal of Network and Computer Applications* 64(2016), 137–153.
51. Zupan, G. (2016a). *E-veščine in digitalna ekonomija*. Pridobljeno 2. maja 2018 iz <http://www.stat.si/StatWeb/File/DocSysFile/8922/e-VESCINE.pdf>
52. Zupan, G. (2016b). *Analiza masovnih podatkov in najemanje storitev računalništva v oblaku v podjetjih z vsaj 10 zaposlenimi*. Pridobljeno 6. maja 2018 iz <http://www.stat.si/StatWeb/News/Index/6170>
53. Zupan, G. (2017). *Ob letošnjem svetovnem dnevu telekomunikacij in informacijske družbe poudarek na masovnih podatkih*. Pridobljeno 11. maja 2018 iz <http://www.stat.si/StatWebArhiv/prikazi-novico?id=6667&idp=25&headerbar=16>

PRILOGE

Priloga 1: Elektronsko sporočilo z vabilom k izpolnitvi spletne ankete

Raziskava stanja uporabe masovnih podatkov v podjetjih - pomoč pri magistrskem delu



Monika Čurin
Za

3. 10. 17



Spoštovani,

sem podiplomska študentka Ekonomske fakultete v Ljubljani, kjer zaključujem študij poslovne informatike.

Na vas se obračam, ker vas vljudno prosim za pomoč pri izvedbi magistrskega dela, z naslovom: Analiza stanja in potencialov uporabe masovnih podatkov za ustvarjanje novih poslovnih priložnosti v slovenskih podjetjih.

V nadaljevanju vam pošiljam povezavo do kratke anonimne spletne ankete, za katero prosim, če jo lahko izpolnite.

Predlagam, da anketo izpolni nekdo iz področja poslovne inteligence, upravljanja podatkovnih baz oziroma nekdo, ki se ukvarja z obdelavo podatkov v vašem podjetju (tehnični, analitični ali drug profil).

Povezava do ankete: <https://www.ika.si/a/140164>

Vnaprej se vam lepo zahvaljujem!

Lep pozdrav,
Monika Čurin

Priloga 2: Anketni vprašalnik: Analiza uporabe masovnih podatkov

Analiza uporabe masovnih podatkov

Kratko ime ankete: Big Data vprašalnik

Dolgo ime ankete: Analiza uporabe masovnih
podatkov

Število vprašanj: 19

Anketa je zaključena.

Aktivna od: 3. 10. 2017

Avtor: monika.curin

Dne: 28. 9. 2017

Opis:

Aktivna do: 3. 1. 2018

Spreminjal: monika.curin

Dne: 12. 11. 2017

Spoštovani, sem podiplomska študentka poslovne informatike na Ekonomski fakulteti v Ljubljani. Pripravljam magistrsko nalogo s področja masovnih podatkov (ang. Big Data). Predvsem me zanimajo trenutno koriščenje masovnih podatkov v velikih slovenskih podjetjih ter priložnosti za dodaten razvoj podjetij na tem področju. Zato vas vljudno prosim za izpolnitev anketnega vprašalnika, ki vam ne bo vzelo več kot 10 minut. Anketa je anonimna, zbrani podatki pa bodo uporabljeni v namen magistrskega dela. Vnaprej se vam lepo zahvaljujem in vas lepo pozdravljam, Monika Čurin

Q1 – Poznate ali ste že slišali za izraz »masovni podatki« (ang. Big Data)?

- ☐ Da.
☐ Ne.

IF (1) Q1 = [2]

Q2 – Izraz masovni podatki (ang. Big Data) predstavlja opis velike količine digitalnih podatkov, ki jih ustvarjamo ljudje z vsakodnevno uporabo različnih spletnih storitev in ki jih je težko upravljati z obstoječimi analitičnimi orodji za obdelavo podatkov. Izraz opisuje tako podatke kot tudi orodja in analitične metode za upravljanje s tako obsežnimi in kompleksnimi podatkovnimi zbirkami. Koncept masovnih podatkov je podoben konceptu poslovne inteligence, oba koncepta opredeljujeta upravljanje podatkov tako z vidika zbiranja, analiziranja, tehnologije in poročanja z namenom pridobivanja novih spoznanj in sprejemanja olajšanih poslovnih odločitev. Razlikujeta se predvsem v količini podatkov, hitrosti pridobivanja podatkov in raznolikosti podatkov. Gre za zbiranje, obdelavo in analiziranje velike količine nestrukturiranih podatkov v realnem času. Menite, da bi lahko bil predstavljeni koncept masovnih podatkov za vaše podjetje zanimiv v prihodnosti?

- ☐ Da.
☐ Ne.

IF (2) Q1 = [1]

Q3 – Ali v vašem podjetju obdelujete t. i. masovne podatke?

- ☐ Da.
☐ Ne, vendar načrtujemo v prihodnosti.
☐ Ne.
☐ Ne vem.

IF (3) Q3 = [1]

Q4 – Prosim, označite, iz katerih od spodaj navedenih virov zbirate in obdelujete podatke v vašem podjetju.

Možnih je več odgovorov.

- ☐ Dokumenti.
☐ Poslovne transakcije in podatkovne baze.
☐ E-pošta.
☐ Slike.
☐ Podatki senzorjev ali drugih pametnih naprav (t. i. IoT-naprav).
☐ Indeksi spletnih strani.

- ☐ Spletni blogi.
- ☐ Družabna omrežja.
- ☐ Telefonski pogovori.
- ☐ Sporočila SMS.
- ☐ Videi.
- ☐ Drugo:

IF (4) Q2 != [1, 2] and Q3 != [2, 3, 4] (Aktivni Big Data)

Q5 – Podatki omogočajo širok nabor različnih metod obdelave podatkov. Zanima nas, ali ste že kdaj uporabili katero izmed spodaj navedenih metod.

	Smo uporabili	Nismo uporabili	Ne vem
Tekstovno rudarjenje (ang.: Text mining)	☐	☐	☐
Napovedna analitika (ang. Predictive analytics)	☐	☐	☐

IF (4) Q2 != [1, 2] and Q3 != [2, 3, 4] (Aktivni Big Data)

Q6 – Ali je v strategiji vašega podjetja zajeto tudi upravljanje masovnih podatkov?

- ☐ Da.
- ☐ Ne, vendar načrtujemo vključitev v prihodnosti.
- ☐ Ne.

IF (4) Q2 != [1, 2] and Q3 != [2, 3, 4] (Aktivni Big Data)

Q7 – Ali merite učinek uporabe masovnih podatkov na rezultate vašega podjetja?

- ☐ Da.
- ☐ Ne, vendar nameravamo v prihodnosti.
- ☐ Ne.

IF (4) Q2 != [1, 2] and Q3 != [2, 3, 4] (Aktivni Big Data)

IF (5) Q7 = [1] (Merijo učinek)

Q8 – Opažate pozitiven ali negativen učinek na poslovanje?

- ☐ Pozitiven.
- ☐ Negativen.

IF (4) Q2 != [1, 2] and Q3 != [1, 2, 3, 4] (Aktivni Big Data)

IF (6) Q7 = [2, 3] or Q8 = [1, 2] (Potencial za vložek)

Q9 – Prosim, če nam lahko zaupate višino vloženih finančnih sredstev (v EUR) vašega podjetja v zadnjih 5 letih v koncept masovnih podatkov.

- ☐ Več kot 5.000.000.
- ☐ 1.000.000–4.999.999.
- ☐ 500.000–999.999.
- ☐ 100.000–499.999.
- ☐ Manj kot 99.999.
- ☐ Finančnih vlaganj ni bilo.
- ☐ Ne vem/Nimam podatka

IF (4) Q2 != [1, 2] and Q3 != [1, 2, 3, 4] (Aktivni Big Data)

IF (7) Q9 = [1, 2, 3, 4, 5] (So vložili)

Q10 – V kateri del vašega podjetja na področju masovnih podatkov ste vložili največ?

- ☐ Vizualizacija podatkov.
- ☐ Zaledni sistemi (na primer: podatkovno skladišče).
- ☐ Orodja za zajem in obdelavo podatkov.
- ☐ Zaposlitev novih sodelavcev z znanjem na področju napredne analitike in masovnih podatkov.
- ☐ Nič od navedenega.

IF (4) Q2 != [1, 2] and Q3 != [1, 2, 3, 4] (Aktivni Big Data)

IF (8) Q9 = [6, 7] or Q10 = [1, 2, 3, 4, 5] (Cilj masovnih podatkov)

Q11 – Na katera področja vašega podjetja poskušate vplivati z aktivnostmi v okviru koncepta masovnih podatkov?

Možnih je več odgovorov.

- ☐ Znižanje operativnih stroškov podjetja.
- ☐ Vzpostavitev t. i. »podatkovne kulture« podjetja (ang. data-driven culture).
- ☐ Vzpostavitev novih poti za razvoj inovacij.
- ☐ Skrajšanje časa na področju razvoja izdelkov in storitev.
- ☐ Ponudba novih izdelkov in storitev.
- ☐ Monetiziranje masovnih podatkov.
- ☐ Preoblikovanje in repozicioniranje podjetja v prihodnosti.
- ☐ Večja konkurenčnost podjetja.
- ☐ Izboljšanje poslovnih procesov.
- ☐ Personalizirano oglaševanje.
- ☐ Izboljšave na področju nujenja pomoči strankam.
- ☐ Drugo:

IF (4) Q2 != [1, 2] and Q3 != [, 2, 3, 4] (Aktivni Big Data)

Q12 – Kako bi ocenili naklonjenost in podporo vašega vodstva aktivnostim na področju masovnih podatkov?

- ☐ Zelo naklonjeni.
- ☐ So naklonjeni.
- ☐ Niti niti.
- ☐ Niso naklonjeni.
- ☐ Sploh niso naklonjeni.

IF (4) Q2 != [1, 2] and Q3 != [, 2, 3, 4] (Aktivni Big Data)

Q13 – Alharthi, Krotov in Bowman ugotavljajo, da je eden od glavnih izzivov organizacij, ki se ukvarjajo z masovnimi podatki, pomanjkanje znanja oziroma podjetjem primanjkuje kadra z ustreznim znanjem na področju analitike (Alharthi in drugi, 2017, str. 288). Se tudi v vašem podjetju soočate s tem izzivom?

- ☐ Da.
- ☐ Ne.

IF (4) Q2 != [1, 2] and Q3 != [, 2, 3, 4] (Aktivni Big Data)

Q14 – Med naštetimi trditvami, prosim, označite najbolj pomemben izziv za vas, s katerim se soočate v okviru masovnih podatkov.

- ☐ Zagotovitev kakovostnih podatkov.
- ☐ Varovanje zasebnosti podatkov.
- ☐ Zagotovitev varnosti podatkov.
- ☐ Zagotovitev finančnih sredstev.
- ☐ Zagotovitev upravljanja s podatki.
- ☐ Drugo:

IF (4) Q2 != [1, 2] and Q3 != [, 2, 3, 4] (Aktivni Big Data)

IF (9) Q5b = [1] and Q14 = [1, 2, 3, 4, 5] (Potencial za BD v IoT)

Q15 – Na začetku vprašalnika ste označili, da imate v vašem podjetju izkušnje z izvedbo napovedi dogodkov oziroma uporabo napovedne analitike. V razmahu je t. i. internet stvari (ang. Internet of Things), ki povezuje različne vrste naprav v internet in tako omogoča zajem vedno večjih količin podatkov. Ali se boste v vašem podjetju v prihodnosti posluževali uporabe napovedne analitike na področju interneta stvari?

- ☐ Da, zagotovo.
- ☐ Da, verjetno.
- ☐ Niti niti.
- ☐ Najverjetneje ne.
- ☐ Zagotovo ne.

C1 – Kalkulacije

BLOK (10) (DEMOGRAFIJA)

Q16 – Izberite panogo, v katero sodi vaše podjetje.

- ☐ Zavarovalništvo.
- ☐ Finance in bančništvo.
- ☐ Telekomunikacije.
- ☐ Logistika.
- ☐ Mediji in založništvo.
- ☐ Komunala.
- ☐ Živilska industrija.
- ☐ Medicina in farmacija.
- ☐ Izobraževanje.
- ☐ Gradbeništvo in nepremičnine.
- ☐ Pravna in odvetniška dejavnost.
- ☐ Gostinstvo.
- ☐ Trgovina na drobno in debelo.
- ☐ Energetika.
- ☐ Drugo:

BLOK (10) (DEMOGRAFIJA)

Q17 – Označite vaš položaj v podjetju.

- ☐ Vodja analitičnega oddelka.
- ☐ Vodja tehničnega oddelka.
- ☐ Vodja oddelka za masovne podatke (ang. Head of Big Data).
- ☐ Član uprave za podatke (ang. Chief Data Officer).
- ☐ Član uprave za trženje (ang. Chief Marketing Officer).
- ☐ Predsednik uprave.
- ☐ Drugo:

BLOK (10) (DEMOGRAFIJA)

XSPOL – Spol:

- ☐ Moški.
- ☐ Ženski.