

Tartu Ülikool
Sotsiaalteaduste valdkond
Ühiskonnateaduste instituut
Sotsiaaltöö ja sotsiaalpoliitika eriala

Katrin Siider

SUURANDMETE KASUTAMISE VÕIMALUSED SOTSIAALTÖÖS: SISSEJUHATAV
KIRJANDUSÜLEVAADE

Magistritöö

Juhendaja: Merle Linno, MSW

Tartu 2021

Olen koostanud töö iseseisvalt. Kõik töös kasutatud teiste autorite tööd, põhimõttelised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

Katrin Siider, 24.05.2021

ABSTRACT

Possibilities to use Big Data in Social Work: A Scoping Review. Katrin Siider (2021)

This thesis aims to give both an overview of the possibilities as well as the challenges to use Big Data in social work.

The development of technology, the spread of the Internet and different digital solutions has resulted in a lot of data being created, collected, stored, analyzed and used every day. These large volumes of data are referred to as Big Data. While Big Data can be used in different fields for different purposes, the field of social work is yet to harness its possible potential.

In this thesis, a scoping review methodology is used to analyze through academic journal articles the possibilities as well as challenges to use Big Data in social work.

The results of this research show that Big Data can be used in social work for four different purposes. First, to develop decision support systems to aid social workers with decision making during different stages of social work practice; second, to conduct social work research; third to enhance social work practice and fourth, to inform social policy development. In doing so, important questions and challenges related to information systems and databases, data quality (including sampling, representativeness, etc.) as well as social work ethics (privacy, informed consent, etc.) must be addressed.

Keywords: Big Data, social work, social policy, social work ethics, decision support systems, information systems, databases

SISUKORD

ABSTRACT	3
SISSEJUHATUS	5
1 UURIMUSE KONTEKST	7
1.1 Suurandmed.....	7
1.2 Sotsiaaltöö infoühiskonnas.....	10
2 METOODIKA.....	14
2.1 Magistritöö metodoloogiline lähtekoht	14
2.1.1 Sissejuhatav kirjandusülevaade (ingl <i>scoping review</i>)	16
2.1.1.1 Uurimusküsimuse sõnastamine	19
2.1.1.2 Asjakohase kirjanduse otsimine	20
2.1.1.3 Allikate valimine	23
2.1.1.4 Andmete kaardistamine	26
2.1.1.5 Materjali läbitöötamine, kokkuvõtte tegemine, tulemuste esitlemine	27
2.1.1.6 Konsultatsioon	29
3 ANALÜÜS	30
3.1 Suurandmete kasutamise võimalused sotsiaaltöös kirjanduse põhjal	30
3.1.1 Otsustustoe süsteemid	30
3.1.2 Sotsiaaltöö praktika uurimine	34
3.1.3 Sotsiaaltöö praktika edendamine	36
3.1.4 Sotsiaalpoliitika kujundamine.....	38
3.2 Suurandmete kasutamisega seotud väljakutsed kirjanduse põhjal	41
4 ARUTELU	44
4.1 Suurandmed sotsiaaltöös	44
4.2 Andmeteadlikkus sotsiaaltöös	48
4.3 Sotsiaaltöö eetika ja suurandmed	50
KOKKUVÕTE	52
KASUTATUD KIRJANDUS	53
5 LISAD	62
5.1 LISA 1. Andmete kaardistamise vorm.....	62

SISSEJUHATUS

Seoses tehnoloogia arenguga, interneti levikuga ning erinevate digitaalsete lahenduste laialdase kasutamisega on märkimisväärselt suurenenud andmete hulk, mida igapäev luuakse, kogutakse, talletatakse, analüüsitakse ja kasutatakse. Sellistele suurtele andmekogustele viidatakse kui suurandmetele (ingl *Big Data*). Suurandmed on n-ö lööksõna, mis Masso jt (2020) hinnangul ei anna edasi suurandmete kui nähtuse mitmetähenduslikkust. Üldistatult saab öelda, et suurandmete tekkeviis ning sellest tulenevalt suurandmetele ligipääs ja andmete kasutamise võimalused on erinevad (Puusalu, 2020).

Suurandmete rakendusvaldkonnad ja rakendusala on mitmekülgsed. Andmeid saab kasutada nii meditsiinivaldkonnas, julgeolekuvaldkonnas, turunduses ja kaubanduses, tööstuses jne (Eljas-Taal, Veerpalu ja Romanainen, 2017). Seda näiteks selleks, et inimeste huvisid, vajadusi ja käitumist paremini mõista, uusi teenuseid välja töötada, tõenditele tuginevaid otsuseid teha, ressursse planeerida ning laiemalt innovatsiooni luua. Coulton jt (2015) kirjutavad, et sotsiaaltöö valdkonnas ei ole suurandmete potentsiaali sotsiaalsete probleemide paremaks mõistmiseks ja neile lahenduste väljatöötamiseks täiel määral veel teadvustatud ja rakendatud. Samas on suurandmetest ja suurandmete analüüsivõimalustest saanud teiste andmeallikate ja andmeanalüüsi meetodite kõrval oluline ressurss sotsiaaltöö praktika ja -poliitika uurimiseks (Gillingham ja Graham, 2017), sotsiaalsetele probleemidele ennetavalt reageerimiseks, sotsiaaltöö valdkonnas uute avastuste tegemiseks, inimeste käitumise ja huvide põhjal sekkumiste täiustamiseks jne (Coulton jt, 2015).

Tööjõuvajaduse seire- ja prognoosisüsteemi OSKA 2020. aastal ilmunud ülevaates valdkonnaspetsiifiliste info- ja kommunikatsioonitehnoloogia oskuste vajaduse kohta kirjutatakse, et järgmise viie aasta kontekstis on sotsiaaltöö valdkonna juhtide, tippspetsialistide ning sotsiaaltöötajate puhul oluline, et nad tunneksid suurandmete kasutamise võimalusi ning oskaksid suurandmeid analüüsida ja analüüsi tulemusi tõlgendada (OSKA ülevaade..., 2020). Siit tekkis huvi, et mida teaduskirjanduses räägitakse suurandmete kasutamise kohta sotsiaaltöös. Kas ja kuidas ning millistel eesmärkidel on suurandmeid sotsiaaltöös kasutatud ja milliseid väljakutseid on sellega seoses tekkinud.

Sellest tulenevalt uurin käesolevas magistritöös, millised on suurandmete kasutamise võimalused ning suurandmete kasutamisega seotud väljakutsed sotsiaaltöö valdkonnas. Nimetatud eesmärkide saavutamiseks püstitasin järgmised uurimisküsimused:

- 1) Millised on suurandmete kasutamise võimalused sotsiaaltöös teadusartiklite põhjal?
- 2) Millised on suurandmete kasutamisega seotud väljakutsed sotsiaaltöös teadusartiklite põhjal?

Uurimisküsimustele vastamiseks viisin läbi sissejuhatava kirjanduse ülevaate. Kuna tehnoloogia, sh suurandmetega seonduv (nt erinevad võimalused suurandmeid koguda, analüüsida ja kasutada) areneb väga kiiresti, siis kasutan oma magistritöö uurimuses teadusartikleid, mis on avaldatud peale 2015. aastat. Magistritöö esimeses peatükis avan uurimuse konteksti, sh kirjutan kuidas erinevad autorid on suurandmeid defineerinud ja liigitanud. Seejärel kirjeldan, kuidas sotsiaaltöö valdkonnas on ajas hoiakutes ja praktikates seoses tehnoloogia kasutamisega toimunud muutus toimunud, mis loob muuhulgas head eeldused suurandmete kasutamisele sotsiaaltöö valdkonnas.

Magistritöö teises peatükis kirjutan magistritöös kasutatud metodoloogilisest lähenemisest. Kõigepealt teen põgusa ülevaate erinevatest kirjandusülevaate liikidest ning seejärel kirjeldan põhjalikult minu poolt valitud sissejuhatava kirjandusülevaate metoodikat ning selgitan, kuidas ja millistele kaalutlustele tuginedes olen erinevaid kirjandusülevaate etappe teostanud.

Kolmandas peatükis analüüsin magistritöö valimi moodustanud teadusartikleid. Artiklite põhjal kirjeldan suurandmete kasutamise võimalusi ja suurandmete kasutamisega seotud väljakutseid sotsiaaltöös. Viimases, neljandas peatükis arutlen magistritöö tulemuste üle.

Soovin tänada enda juhendajat Merle Linnot, kes oli toetav ja abivalmis ning andis magistritöö kirjutamisel häid soovitusi. Samuti tänan oma retsensenti Maris Männistet heade tähelepanekute ja ettepanekute eest.

1 UURIMUSE KONTEKST

1.1 Suurandmed

Suurandmed on n-ö lööksõna, millel puudub üks kindel definitsioon ning käsitlus. Eesti keele Instituudi mitmekeelses terminibaasis (2014) on suurandmeid defineeritud kui "suured mahud struktureeritud ja struktureerimata andmeid, mille haldamine tavapäraste relatsioonilise andmebaasi- ja andmetöötluse vahenditega on raskendatud, kui mitte võimatu (maht, formaadid, kiirus)". Uuringu „Lingitud Eesti“ (2014) lõpparuandes defineeritakse suurandmeid kui andmeid, mille „maht on nii suur, et traditsiooniliste andmebaasisüsteemide võimekus pole piisav nende haldamiseks“. Eesti Arenguseire Keskuse (2018) kohaselt on suurandmed seevastu „suured ja keerulised andmekogumid, mida on olemasolevate infotehnoloogiliste rakendustega keeruline analüüsida“. Nagu näha, siis definitsioonides rõhutatakse suurandmete mahtu ja seda, et suurandmete haldamiseks ja analüüsimiseks vaja vastavat tehnilist võimekust.

Kuigi suurandmetel puudub n-ö universaalne definitsioon, siis kõige laialdasemalt kasutatakse suurandmete kirjeldamisel D. Laney poolt 2001. aastal sõnastatud suurandmete omadusi, milleks on maht, kiirus ja mitmekesisus (2012). Inglise keeles algavad nimetatud omadused v-tähega (maht, ingl *volume*; kiirus, ingl *velocity*; mitmekesisus, ingl *variety*), mistõttu räägitakse suurandmete kirjeldamisel tihti suurandmete v-dest (antud juhul ingl *The 3 V's of Big Data*). Maht viitab sealjuures andmestiku suurusle, st andmeid mõõdetakse terabaitides ja petabaitides, mistõttu on suurandmete analüüsimiseks vaja uusi andmetöötlusvorme; mitmekesisus tähendab, et suurandmete puhul on tegemist erinevat tüüpi andmetega (nt pildid, videod, sotsiaalmeedia postitused); ning kiirus viitab andmete tekkimise, andmeedastuse ning analüüsimise ja töötlemise kiirusele (Gandomi ja Heider, 2014; Günther, Rezazade, Huysman ja Feldberg, 2017; Puusalu, 2020; Küngas jt, 2014; McKinsey Global Institute, 2011).

Ülal väljatoodud D. Laney poolt sõnastatud kolmele v-le lisaks on aja jooksul tähelepanu juhitud veel teistele suurandmeid kirjeldavatele mõõtmetele. Näiteks räägitakse suurandmete tõepärasusest (ingl *veracity*), mis hõlmab endas andmete kvaliteedi ja usaldusväärsusega seotud aspekte (Altintas ja Gupta, 2020; Gandomi ja Heider, 2014; Hilbert, M. 2020; Ishwarappa ja Anuradha, 2015; Küngas jt, 2014; Santiago ja Smith, 2019), andmete väärtusest (ingl *value*) ehk suurandmete võimalikust kasutegurist (Altintas ja Gupta, 2020; Gandomi ja

Heider, 2014; Ishwarappa ja Anuradha, 2015; Zetino ja Mendoza, 2019; Santiago ja Smith, 2019) ning andmete keerukusest (ingl *variability*) ehk sellest, et andmed pärinevad erinevatest allikatest (Gandomi ja Heider, 2014).

Baur jt (2020) lisavad, et kuigi suurandmed on rohkem tähelepanu pälvinud viimase paarikümne aasta jooksul, siis ei ole suurandmete puhul tegemist n-ö uue nähtusega, vaid et tehnoloogia areng on andmete loomisse, kogumisse, talletamisse, analüüsimisse ja kasutamisse uue mõõtme toonud. Nad eristavad n-ö traditsioonilisi suurandmeid (administratiivsed andmed, millest on saanud digitaalsed andmed, st andmed on digitaalseks muudetud või andmeid on hakatud digitaalselt koguma) ja n-ö uut tüüpi suurandmeid (andmed, mis on veeb 2.0 tehnoloogiate kasutamise kõrvalprodukt, nt sotsiaalmeedia kasutamisel, veebilehtede külastamisel, nutikellade kandmisel jne tekkivad andmed).

Sarnaselt, Puusalu (2020) kirjutab, et suurandmed moodustuvad andmekogudes ja andmebaasides talletatud andmetest ning interneti rakenduste (nt sotsiaalmeediakanalite) ja nutistu seadmete (nt nutitelefonide, nutikellade) kasutamisel tekkivatest andmetest. Ta lisab, et andmeid, mida talletatakse erinevates andmekogudes ja andmebaasides kogutakse tavaliselt kindlal eesmärgil, need on seotud konkreetsete teenuste või süsteemidega ning andmete kogumiseks on määratud kriteeriumid ja võimalik et seadusandlik raamistik. Ühtlasi leiab Puusalu (2020), et inimesed on andmebaasi olemasolust ja sellest et nende andmeid andmebaasi kantakse tavaliselt teadlikud. Interneti rakenduste kasutamisel seevastu tekivad nii sellised andmed, mille kogumisest on inimesed teadlikud kui ka sellised andmed, mille kogumisest inimesed teadlikud ei ole (Puusalu, 2020). Puusalu (2020) selgitab, et kui näiteks e-poest oste sooritades inimesed üldjuhul teavad, et selle tulemusel tekivad andmed, siis sotsiaalmeedias postitusi tehes, teiste inimeste postitusi meeldivaks märkides jne ei pruugi inimesed teadvustada, et seda tehes loovad nad andmeid, mida kogutakse ja analüüsitakse. Samuti ei ole inimesed üldjuhul teadlikud metaandmete kogumisest, mis interneti rakenduste kasutamisel tekivad (nt külastatud veebilehekülgede arv ja olemus). Ka nutistu seadmete kasutamise tulemusel tekivad andmed, mida nagu Puusalu (2020) kirjutab on võimalik teistesse seadmetesse ja andmebaasidesse saata ning seejärel analüüsida.

Nagu näha, siis suurandmeid on võimalik erinevat moodi defineerida, kirjeldada ja liigitada. Erinevate autorite seas puudub üksmeel selles osas, mis täpselt eristab suurandmeid teistest andmetest. Kui näiteks Kitchen ja McArdle (2016) leiavad, et just andmete maht ja andmete tekkimise kiirus on suurandmeid teistest andmetest eristavad omadused, siis näiteks Boyd ja

Crawford (2012) leiavad, et suurandmete puhul ei ole mitte andmete maht määrava tähtsusega, vaid võimekus suuri andmestikke kokku koondada ja ristviidata. Kirt (2020) kirjutab, et erinevad organisatsioonid võivad suurandmeid erinevalt määratleda tulenevalt organisatsiooni andmete töötlemise tehnilisest võimekusest. Ta lisab, et suurandmed võivad ajas koos tehnoloogia arenguga muutuda. See tähendab andmed, mida täna käsitletakse suurandmetena ei pruugi tulevikus seda enam olla.

Elevus, mis suurandmeid ümbritseb on paljuski seotud ülal kirjeldatud suurandmete omadustega. Suur hulk mitmekesiseid andmeid, mida tekib koguaeg juurde ja mida on võimalik kiiresti töödelda ja analüüsida – see on kõikides valdkondades suurandmete vastu huvi äratanud. Samas leiavad erinevad autorid (Boyd ja Crawford, 2012; Crawford, Miltner ja Gray, 2014), et nimetatud omadusi ei tohi ületähtsustada.

Näiteks Crawford jt (2014) kirjutavad, et kuigi suurandmete fenomeni seostatakse tihti just tehnoloogia arenguga (st tänu tehnoloogia arengule on võimalik rohkem andmeid koguda, talletada, analüüsida, omavahel seostada jne), siis on suurandmete kui nähtuse mõtestamisel oluline arvestada laiemat kultuurilist, sotsiaalset, poliitilist ja majanduslikku kontekstiga. Nad kirjeldavad, kuidas suurandmeid käsitlevas kirjanduses ülistatakse suurandmete potentsiaali innovatsiooni loomisel, andmetest seoste leidmisel, inimeste käitumise ennustamisel, personaalsete teenuste loomisel jne. Andmed justkui annavad tõese ja objektiivse ülevaate maailmast ja nähtustest, mida andmete põhjal kirjeldada soovitakse (Crawford jt, 2014; Boyd ja Crawford, 2012). Crawford jt (2014) leiavad, et just see on suurandmed ahvatlevaks teinud. Masso, Vihalemm ja Saarniit (2020), viidates Jones-Devitt ja Samiei (2011) ja (Peters, 2001) kirjutavad, et n-ö iseenda eest rääkivate andmete põhjal tehtud otsuseid peetakse täpsemaks, ratsionaalsemaks ja süsteemsemaks, sest kõik muu, millele inimesed otsuste tegemisel toetuvad (nt kogemused, emotsioonid, isiklik otsustamisstiil) jääb kas täielikult või osaliselt n-ö tahaplaanile.

Samas, Crawford jt (2014) hinnangul on oluline, et andmete analüüsimisel arvestatakse kontekstiga, kus andmed on loodud. Andmeid võib olla küll palju aga see ei taga ilmtingimata andmete esinduslikkust (Crawford jt, 2014; Boyd ja Crawford, 2012; Baur jt, 2020). Lisaks sellele on oluline teadvustada, et andmete loomisel, analüüsimisel ning kasutamisel tehakse subjektiivseid valikuid, seda konkreetsetes valdkondlikes, institutsionaalses aga ka laiemas ühiskondlikes kontekstis (Boyd ja Crawford, 2012). Seega on andmete tehnilise analüüsi

kõrval oluline, milliseid tähendusi andmete põhjal luuakse ja kuidas andmeid tõlgendatakse (Latzer jt, 2018; viidatud Masso jt, 2020).

Olenemata sellest, kuidas suurandmeid defineerida, kirjeldada või liigitada ei ole suurandmed n-ö väärtus iseenesest, vaid andmetest on vaja väärtust luua. Suurandmete analüüsimiseks on erinevaid võimalusi. Kasutan sarnaselt Reddenile, Denickule ja Warnele (2020) oma magistritöös läbivalt terminit andmetehnoloogiad (mis hõlmab masinõppe meetodite, algoritmide jms rakendamist andmetele, et andmeid analüüsida, andmetest mustreid ja seoseid leida jne).

1.2 Sotsiaaltöö infoühiskonnas

Infotehnoloogia kiire areng on oluliselt muutnud ühiskonda. Võimaluste kõrval (nt paremad tooted ja teenused, parem ligipääs infole) on mitmeid väljakutsed (nt digitaalne lõhe, erinevused oskustes). Goldkind ja Wolf (2015) kirjutavad, et kuigi tehnoloogia areng muudab pidevalt võimaluste maastikku, siis ei garanteeri see suuremat sotsiaalset õiglust. Infotehnoloogia areng ei tähenda seega, et kõik inimesed on saanud või saavad võrdselt osa sellega kaasnevast majanduslikust ja sotsiaalsest arengust ning heaolust. Sotsiaaltööl ning sotsiaaltöötajatel on ühiskonnas toimuvate muutuste kontekstis oluline roll sotsiaalse õigluse edendamisel ning inimeste sotsiaalse heaolu toetamisel.

Sotsiaaltöö valdkond on tehnoloogia pakutavate võimaluste rakendamisel sotsiaaltöös olnud pigem aeglane ja ettevaatlik (Berzin, Singer ja Chan, 2015; Parrot ja Madoc-Jones, 2008; Goldkind, Wolf ja Freddolino, 2020). Seda erinevatel põhjustel. Piiratud ressursside, väheste oskuste ja erinevate eetiliste küsimuste kõrval on sotsiaaltöö valdkonnas kardetud, et tehnoloogia kasutuselevõtmine sotsiaaltöös muudab sotsiaaltööd, sh vähendab võimalusi valdkonnas ajalooliselt oluliseks peetud inimeste jutustuste talletamiseks ning näost-näku suhtlemiseks (Goldkind jt, 2020). Samas, vaadates sotsiaaltöö valdkonna standardeid ning eetikakoodekseid on näha muutust (Goldkind jt, 2020; Educational..., 2015; NASW, 2021; Reamer, 2013; Eetikapõhimõtted..., 2018). Hirm ja kahtlused tehnoloogia kasutamise osas on

asendunud teadvustamisega, et tehnoloogia oluliselt mõjutab sotsiaaltöö praktikat ning pakub uusi võimalusi sotsiaaltöö tegemiseks.

Goldkind, Wolf ja Freddolino (2020) kirjeldavad, kuidas Ameerika Ühendriikide hariduspoliitika ja akrediteerimise standardites (ingl *The Educational Policy and Accreditation Standards*) on tehnoloogia kasutamisega seonduvat sotsiaaltöötajate professionaalses ettevalmistuses käsitletud. Nad võrdlevad 2001., 2008. ja 2015. aasta standardeid ja kirjeldavad, kuidas tehnoloogia käsitlemine neis on muutunud. Kui 2001. ja 2008. aasta standardid sisaldasid üksnes põgusaid viiteid tehnoloogiale (Goldkind jt, 2020), siis 2015. aasta standardis on sotsiaaltöötajate kompetentside kirjeldamisel tehnoloogia juba rohkem esile tõstetud (Educational..., 2015; Goldkind jt, 2010). Täpsemalt, seoses sotsiaaltöötaja eetilist ja professionaalset käitumist kirjeldava kompetentsiga on välja toodud, et sotsiaaltöötajad peavad tehnoloogiat kasutama eetiliselt ja asjakohaselt, et praktikat toetada (Educational..., 2015). Lisaks sellele on standardis kirjas, et ka õppetöös (sh praktika läbiviimisel) võib erinevaid tehnoloogiaid kasutada (Educational..., 2015; Goldkind jt, 2020).

Tehnoloogia käsitus on muutunud ka Ameerika Ühendriikide sotsiaaltöötajate organisatsioonide (ingl *The National Association of Social Workers, NASW, Association of Social Work Boards, ASWB, Council on Social Work Education, Clinical Social Work Association*) koostöös 2005. aastal valminud ja 2017. aastal täiendatud tehnoloogia ja sotsiaaltöö praktika standardites (ingl *Standards for Technology and Social Work Practice*) (Goldkind jt, 2020). Goldkind jt (2020) selgitavad, et kui 2005. aasta standardites oli rõhuasetus tehnoloogia kasutamisega seotud ohtudel ja riskidel (rõhutati, et tehnoloogia kasutamisel sotsiaaltöös väheneb sotsiaaltöötaja ja klientide näost-näku suhtlemine), siis 2017. aasta standardites rõhutatakse juba tehnoloogia kasutamisega seotud võimalusi. Sellest tulenevalt on täiendatud standardite eesmärk toetada sotsiaaltöötajat tehnoloogia eetilisel kasutamisel sotsiaaltöös ning ühtlasi teavitada laiemat avalikkust sellest, millistele põhimõtetele tuginedes sotsiaaltöötajad oma töös tehnoloogiat kasutavad (Standards..., 2017).

Sarnaselt sotsiaaltöö standarditele on ka valdkondlikud eetikakoodeksid ajas muutunud. Ameerika Ühendriikide sotsiaaltöö assotsiatsiooni eetikakoodeks (ingl *The National Association of Social Workers Code of Ethics*) võeti esimest korda vastu 1960. aastal (NASW, 2021). Alates sellest ajast on eetikakoodeksit mitmel korral muudetud ning esimest korda lisati eetikakoodeksisse viited tehnoloogia kasutamise kohta (täpsemalt elektroonilise meedia kasutamise kohta) 1996. aasta muudatustega, millega pöörati tähelepanu teadliku nõusoleku,

privaatsuse ja konfidentsiaalsusega seotud küsimustele elektroonilise meedia kasutamisel klientidele teenuste pakkumisel (Reamer, 2013; NASW, 2021). Viimased muudatused eetikakoodeksisse on tehtud 2020. aasta novembris ja 2021. aasta veebruaris ning enamik neist on seotud tehnoloogia kasutamisega sotsiaaltöös (NASW, 2021).

Eesti sotsiaalala töötaja eetikakoodeksis, mis on vastu Eesti sotsiaaltöö III kongressil 30. novembril 2005. aastal (Sotsiaalala..., 2005) viited erinevate tehnoloogiate kasutamisega seotud eetilistele küsimustele puuduvad, kuid Rahvusvahelise Sotsiaaltöö Kõrgkoolide Assotsiatsiooni (ingl *International Association of Schools of Social Work*) eestvedamisel loodud dokumendis „Eetikapõhimõtted sotsiaaltöös“ (ingl *Global Social Work Statement of Ethical Principles*, 2018), milles sätestatud järgitakse ka Eestis, juba kirjeldatakse digitehnoloogia ja sotsiaalmeedia eetilise kasutamise põhimõtteid.

Muutused standardites ja eetikakoodeksites on kahtlemata tingitud sotsiaaltöö praktika reaalsusest. Teaduskirjanduses kirjeldatakse, kuidas tehnoloogia kasutamine sotsiaaltöös on loonud uusi võimalusi sotsiaaltöö teenuste pakkumiseks ning teenustele ligipääsu tagamiseks, koostööks, huvikaitseks, kogukondade aktiveerimiseks, osaluseks jne. Sealjuures on oluline, et infotehnoloogia kasutamine sotsiaaltöös ei asenda sotsiaaltöö valdkonnas ajalooliselt olulist näost-näku kontakti ja suhet inimestega, vaid toetab traditsioonilisi lähenemisi ja praktikaid (Berzin, Singer ja Chan, 2015).

Tehnoloogia võimaldab inimestele teenuseid pakkuda nii telefoni, video kui sotsiaalmeedia vahendusel kasutades muuhulgas mängustamist ning virtuaalreaalsust (Barak jt, 2009; Kumar jt, 2013; viidatud Berzin jt, 2015). Seeläbi on võimalik sotsiaaltöö teenustele ligipääsu tagada just sellistele sihtrühmadele, kellel see varasemalt erinevatel põhjustel raske oli, nt maapiirkonnas ja ääremaal elavatele inimestele (Barsky, 2017; Parrott ja Madoc-Jones, 2008; Steiner, 2020) ning erivajadustega inimestele (Parrott ja Madoc-Jones, 2008; Steiner, 2020). Steiner (2020) selgitab, et n-ö madala lävepaku teenused nagu video vahendusel nõustamine motiveerib inimesi sotsiaaltöötajate poole pöörduma. Sotsiaalmeediakanalite kasutamine omakorda võib olla hea võimalus just noortega kontakti saada (Chan ja Ngai, 2019; viidatud Steiner, 2020).

Lisaks sellele saavad sotsiaaltöötajad erinevaid info ja kommunikatsioonitehnoloogia (IKT) võimalusi kasutades nii ise võrgustikutööd teha (Barsky, 2017) kui ka aidata oma klientidel teiste inimestega võrgustada (nt veebipõhiste tugigruppidega liituda vms) (Parrott ja Madoc-Jones, 2008). Sarnaselt ka huvikaitsega, kus sotsiaaltöötajad saavad IKT vahendeid kasutades

kogukondi aktiveerida (Barksy, 2017) või siis toetada kliente, et nemad ise saaksid organiseeruda ning oma huvide ja õiguste eest seista (Parrott ja Madoc-Jones, 2008). Virtuaalsed sotsiaaltöötajate kogukonnad omakorda võimaldavad sotsiaaltöötajatel muuhulgas infot vahetada ja häid praktikaid jagada ning seeläbi teenuseid täiustada (Adedoyin & Christson, 2016; viidatud Parrott ja Madoc-Jones, 2008).

Goldkind ja Wolf (2015) kirjutavad, et on neli tehnoloogiat, mis n-ö raputavad sotsiaaltööd ja millega sotsiaaltöö valdkond peab kohanema. Nendeks on asjade internet (ingl *The Internet of Things*), mängustamine ehk mänguelementide kasutamine sotsiaaltöös (ingl *gamification*), mobiilne tehnoloogia ja suurandmed. Tõepoolest, nagu ülal selgus, siis esimesed kolm on sotsiaaltöö lähenemistesse ja praktikatesse juba jõudnud. Palju vähem on aga teada suurandmete kasutamise võimaluste kohta sotsiaaltöös.

Andmetehnoloogiate kasutamine sotsiaaltöö valdkonnas võimaldaks erinevaid andmeid omavahel ühendada ning seeläbi keerulisi sotsiaalseid probleeme uurida, teenuseid täiustada ja selle tulemusel inimeste elus ning kogukondades positiivset muutust luua (Santiago ja Smith, 2019). Lisaks sellele võimaldaks erinevate andmeallikate ühendamise inimeste ja kogukondade vajadustele ennetavalt reageerida, teenuste vajadust ennustada ning seeläbi sotsiaaltöö teenuseid pakkuvatel organisatsioonidel oma ressursse tõhusamalt planeerida ja suunata (Walker ja Fisman, 2015; viidatud Santiago ja Smith, 2019).

Nagu öeldud, siis on vähe teada suurandmete kasutamise kohta sotsiaaltöös. Samas, Eesti tööjõuvajaduse seire- ja prognoosisüsteemi OSKA ülevaate kohaselt peavad sotsiaaltöötajad suurandmete kasutamise võimalusi tundma ning lisaks sellele oskama suurandmeid analüüsida ja analüüsitulemusi tõlgendada. Seega tööturu kontekstis on suurandmete alased teadmised olulised.

Probleemipüstitusest lähtuvalt on magistritöö eesmärk uurida, millised on suurandmete kasutamise võimalused ja suurandmete kasutamisega seotud väljakutsed sotsiaaltöös. Selle eesmärgi saavutamiseks olen püstitanud järgmised uurimisküsimused:

1. Millised on suurandmete kasutamise võimalused sotsiaaltöös teadusartiklite põhjal?
2. Millised on suurandmete kasutamisega seotud väljakutsed sotsiaaltöös teadusartiklite põhjal?

2 METOODIKA

2.1 Magistritöö metodoloogiline lähtekoht

Kirjanduse analüüsimine on akadeemilise uurimuse loomulik osa, mis aitab uurimusvaldkonda kaardistada ning uurimuse eesmärgi ja uurimusküsimusi sõnastada (Xiao ja Watson, 2019; Snyder, 2019; Sutton jt, 2019). Kirjandusülevaated võib tinglikult kaheks jagada:

- 1) kirjandusülevaate tegemine kui n-ö eeltöö enne suurema uurimuse läbiviimist, et ennast uuritava teemaga põhjalikult kurssi viia ning olemasoleva kirjanduse põhjal täiendava uurimuse läbiviimise vajadust hinnata ja põhjendada;
- 2) kirjandusülevaate tegemine kui eraldiseisva uurimuse läbiviimine, et olemasoleva kirjanduse analüüsimise tulemusel konkreetsetele uurimusküsimustele vastata (Aveyard, 2014; Xiao ja Watson, 2019).

Kirjandusülevaadet kasutatakse uurimusmeetodina erinevatel eesmärkidel, milleks võib muuhulgas olla teooriate analüüsimine ja/või edasi arendamine, hüpoteeside testimine, valdkonnas olemasoleva teadmuse analüüsimine, kindla teema või kindla probleemi kohta ülevaate andmine jne (Snyder, 2019; Xiao ja Watson, 2019). Sealjuures peab uurija, tulenevalt oma uurimuse eesmärgist, valima sobiva kirjandusülevaate meetodi. Nimetatud meetodite arv on ajas kasvanud (Grant ja Booth, 2009; Sutton, 2019). Selle põhjuseks peetakse tõenduspõhise praktika levikut, mis on muuhulgas kaasa toonud vajaduse olemasolevast teadmusest kokkuvõtteid teha (arvestades sealjuures erinevate valdkondade vajaduste, otsuste tegemiseks olemasoleva aja jms), et seeläbi toetada paremate, tõenditele põhinevate otsuste tegemist (Arksey ja O'Malley, 2005; Grant ja Booth, 2009; Peters jt, 2015).

Grant ja Booth avaldasid 2009. aastal artikli, kus nad andsid ülevaate 14 erinevast kirjandusülevaate meetodist, kuid nagu öeldud, siis meetodeid on aja jooksul juurde tulnud. Nii tuvastasidki Sutton jt oma 2019. aasta uurimuses juba 48 erinevat kirjandusülevaate liiki, mille nad jagasid ülevaate eesmärgi ja tunnuste põhjal seitsmesse gruppi:

- 1) traditsioonilised kirjandusülevaated, nt kriitiline kirjandusülevaade (ingl *critical review*), narratiivne kirjandusülevaade (ingl *narrative review*);
- 2) süstemaatilised kirjandusülevaated, nt meta-analüüs (ingl *meta-analysis*), *Cochrane'i* kirjandusülevaated sekkumiste efektiivsuse kohta (ingl *Cochrane review of effects*);

- 3) kirjandusülevaated kirjandusülevaadete kohta, nt ülevaade (ingl *overview*), katusülevaade (ingl *umbrella review*);
- 4) kiired ülevaated, nt kiire tõendite hindamine (ingl *rapid evidence assessment*);
- 5) kvalitatiivsed kirjandusülevaated, nt kvalitatiivne tõendite süntees (ingl *qualitative evidence synthesis*), kvalitatiivne meta-süntees (ingl *qualitative meta-synthesis*);
- 6) kombineeritud meetodil põhinevad kirjandusülevaated, nt Bayesi meta-analüüs/Bayesi lähenemine (ingl *Bayesian Meta-Analysis/Bayesian Approach*), narratiiv süntees (ingl *narrative synthesis*);
- 7) eesmärgi spetsiifilised kirjandusülevaated, nt sissejuhatav kirjandusülevaade (ingl *scoping review*), metodoloogiline kirjandusülevaade (ingl *methodological review*).

Nii Grant ja Booth (2009) kui Sutton (2019) tõdevad, et erinevate meetodite teineteisest eristamine ei ole alati kerge, terminoloogia on mitmekesine ja erinevate ülevaate liikide puhul kohati kattuv ning meetodid ja kirjandusülevaate läbiviimise protsess ei ole alati selge jne.

Snyder (2019) kirjutab, et sõltumata meetodist sisaldab kirjandusülevaade alati kirjandusülevaate planeerimist (st uurimuse eesmärgi ja uurimusküsimuste sõnastamist, kirjanduse otsinguprotsessi läbi mõtlemist ja planeerimist, sh allikate valikukriteeriumite määratlemist), kirjandusülevaate läbiviimist (st otsinguprotsessi tulemusel tuvastatud allikate analüüsimist), tulemuste analüüsimist (st sellise info kogumist allikatest, mis võimaldab uurimusküsimustele vastata) ja kirjandusülevaate kirjutamist (st analüüsi tulemustest ülevaate andmist). Xiao ja Watson (2019) räägivad sarnastest etappidest selle erinevusega, et nemad käsitlevad tulemuste analüüsimist kirjandusülevaate läbiviimise etapi osana. Kõik kolm autorit aga rõhutavad, et kogu kirjandusülevaate läbiviimise protsess peab olema läbipaistev. Just kirjandusülevaate läbiviimise protsessi, st metodoloogilist lähenemist analüüsides on võimalik tuvastada, millisel moel erinevad ülevaate liigid teineteisest erinevad (Grant ja Booth, 2009; Sutton jt, 2019). Kui Grant ja Booth (2009) leidsid, et on vähe selliseid kirjandusülevaate liike, mille läbiviimiseks on konkreetne metodoloogiline lähenemine olemas, siis kümme aastat hiljem leidsid Sutton jt (2019) oma uuringus, et rohkem või vähem konkreetset suunised on loodud ja et need saab jagada formaalseteks juhisteks, metodoloogilisteks nõuanneteks ja n-ö praktika raames välja kujunenud lähenemisteks.

Vaatamata rohkem või vähem konkreetsete suuniste olemasolule seisavad uurijad kirjandusülevaadet tehes siiski silmitsi erinevate väljakutsetega. Selleks võib olla sobiva metodoloogilise lähenemise valimine või allikate valikukriteeriumite sõnastamine aga ka

segadus selles osas, milliseid andmeid valitud allikatest koguda (Snyder, 2019). Sarnaseid tähelepanekuid tegid ka Sutton jt (2019), kes kirjutavad, et kui näiteks allikate otsimisprotsessi kaardistamiseks on erinevaid instrumente loodud (nt PRISMA diagramm, mida kasutatakse süstemaatiliste kirjandusülevaadete puhul ja mille eesmärk on tagada otsinguprotsessi läbipaistvus ja korratavus), siis leitud allikatest andmete kogumiseks juhised jällegi puuduvad. Lisaks sellele kirjutavad nad, et on kirjandusülevaateid (nt eesmärgi spetsiifilised kirjandusülevaated), mille läbiviimisel soovitatakse erinevaid tegevusi (nt asjakohase kirjanduse otsimise protsessi) n-ö jooksvalt üle vaadata, täiendada ja/või vastavalt vajadusele muuta. Samas, nagu nad väljatoovad, siis puuduvad konkreetsed juhised, kuidas seda täpselt teha. Samuti puuduvad juhised selleks, kuna allikate otsimine lõpetada (Sutton jt, 2019). Seega saab öelda, et sõltuvalt kirjandusülevaate liigist võib tekkida olukord, kus uurija peab kirjandusülevaate erinevates etappides ise otsustama, kuidas vajalikke tegevusi läbi viia (Sutton jt, 2019).

Kokkuvõtvalt, kirjandusülevaateid tehakse erinevates valdkondades, erinevatel põhjustel ja erinevatele uurimusküsimustele vastamiseks (Munn jt, 2019). Saab öelda, et ei ole ühte ja n-ö ainuõiget kirjandusülevaate liiki. Metoodika valik sõltub uurimuse eesmärgist ja uurimusküsimustest.

2.1.1 Sissejuhatav kirjandusülevaade (ingl *scoping review*)

Käesolevas uurimustöös kasutan sissejuhatava kirjandusülevaate metoodikat, mille läbiviimiseks sõnastasid esimest korda suunised Arksey ja O'Malley (2005). Nimetatud lähenemise valimisel lähtusin järgnevast:

- 1) esiteks, sissejuhatava kirjandusülevaate metoodikat peetakse sobivaks just sellistes valdkondades ja sellistel teemadel kirjandusülevaate tegemiseks, mida ei ole varem põhjalikult analüüsitud, mis on alles esile kerkimas või kus teadmus on alles kujunemas (Levac jt, 2010; Peters jt, 2015; Tricco jt, 2016) Munn jt, 2019).
- 2) teiseks, sissejuhatav kirjandusülevaade on asjakohane valik, kui uurimuse eesmärk on anda ülevaade konkreetsest teemast, sh sellest, kui palju ja missugust kirjandust konkreetsetes teemavaldkonnas on üldse avaldatud ja millele see keskendub (Grant ja Booth, 2009; Munn, 2019; Sutton jt, 2019).

- 3) kolmandaks, sissejuhatava kirjandusülevaate läbiviimisel on lubatud kogu uurimuse vältel uurimuse erinevates etappides vastavalt vajadusele täiendusi ja muudatusi teha (Arksey ja O'Malley, 2005; Levac jt, 2010; Peters jt, 2015, 2017, 2020).

Arksey ja O'Malley (2005: 21) kirjeldavad sissejuhatavat kirjandusülevaadet kui ülevaadet, mille „eesmärk on kiiresti kaardistada uurimisvaldkonna põhimõisteid ja peamisi teadmuse allikaid ning liike ja mida võib teostada kui eraldiseisvat projekti, seda eriti valdkonnas, mis on keeruline ja mida ei ole varasemalt põhjalikult uuritud“. Nad selgitavad, et sissejuhatavat kirjandusülevaadet võib teha selleks et: 1) uurida kui palju ja millist kirjandust konkreetsetes teemavaldkonnas üldse on, 2) otsustada, kas süstemaatilise kirjandusülevaate tegemine on eesmärgipärane, 3) teha kokkuvõtteid uurimistulemustest ning jagada uurimistulemusi või 4) leida lünki/tühimikku olemasolevas kirjanduses.

Nagu öeldud, siis Arksey ja O'Malley (2005) olid esimesed, kes pakkusid välja metodoloogilise lähenemise sissejuhatava kirjandusülevaate läbiviimiseks. Nende ettepaneku kohaselt on sellisel ülevaatel viis etappi: 1) uurimusküsimuste sõnastamine, 2) asjakohase kirjanduse otsimine, 3) allikate valimine, 4) andmete kaardistamine, 5) materjali läbitöötamine, kokkuvõtte tegemine ja tulemuste esitamine ning 5) konsultatsiooni läbiviimine (valikuline etapp). Ühtlasi tegid nad oma artiklis üleskutse jätkata diskussiooni sissejuhatava kirjandusülevaate metodoloogilise lähenemise osas. Kirjanduse põhjal (Levac jt, 2010; Daudt jt, 2013; Tricco jt, 2014, 2016; Colquhoun jt, 2014; Peters jt, 2015, 2017, 2020; Khalil jt, 2016; O'Brien jt, 2016; Munn jt, 2018; Teare ja Taks, 2020) saab öelda, et käib elav arutelu nii sissejuhatava kirjandusülevaate terminoloogia, definitsiooni kui metodoloogilise lähenemise osas.

Arksey ja O'Malley kasutasid oma 2005. aasta artiklis terminit sissejuhatav uurimus (ingl *scoping study*), kuid kirjanduses kasutatakse ka teisi termineid, nt sissejuhatav kirjandusülevaade (ingl *scoping review*), sissejuhatav meetod (ingl *scoping method*), süstemaatiline kaardistamine (ingl *systematic mapping*), tõendite kaardistamine (ingl *evidence mapping*), kirjanduse kaardistamine (ingl *literature mapping*) jne (Levac jt, 2010; Colquhoun, 2014; Tricco jt, 2014; Pham jt, 2014; O'Brien jt, 2016). Pham jt (2016) leidsid oma uuringus, et autorid, kelle töö rõhuasetus on kirjandusest ülevaate andmisel eelistavad kasutada terminit sissejuhatav kirjandusülevaade ja autorid, kelle töö rõhuasetus on kirjandusülevaate tulemuste analüüsimisel eelistavad terminit sissejuhatav uurimus. Vaatamata terminoloogia mitmekesisusele on siiski leitud, et sellist tüüpi kirjandusülevaateid teinud uurijad kasutavad

kõige rohkem terminit sissejuhatav kirjandusülevaade (Tricco jt, 2014; Pham jt, 2014). Saab öelda, et arutelu sellel teemal ilmselt jätkub, sest uurijate seas puudub üksmeel selles osas, milline termin on kõige paremini kooskõlas seda tüüpi kirjandusülevaate olemuse, eesmärkide ja metodoloogilise lähenemisega. Käesolevas uurimuses kasutan läbivalt terminit sissejuhatav kirjandusülevaade.

Sarnaselt terminoloogiale käib arutelu sissejuhatava kirjandusülevaate definitsiooni üle. Tricco jt (2014) ning Pham jt (2016), kes on uurinud, kuidas on erinevaid sissejuhatavaid kirjandusülevaateid läbi viidud kirjutavad, et enamasti defineeritakse sissejuhatavat kirjandusülevaadet kui ülevaadet, mille eesmärk on olemasoleva kirjanduse kaardistamine, analüüsimine või kirjeldamine. Sõltumata definitsioonist rõhutatakse, et ülevaade peab olema läbiviidud põhjalikult ja läbipaistvalt ning et kogu kirjandusülevaate läbiviimise protsess peab olema korratav (Arksey ja O'Malley, 2005; Levac jt, 2010; Peters jt, 2015, 2017, 2020).

Arksey ja O'Malley (2005) olid küll esimesed, kes pakkusid välja metodoloogilise lähenemise sissejuhatava kirjandusülevaate tegemiseks, kuid nagu ülal väljatoodud, siis nende artikli eesmärk oli muuhulgas algatada arutelu väljapakutud metodoloogia üle. Erinevad uurijad (Levac jt, 2010; Doudt jt, 2013; Peters jt, 2015, 2017, 2020; Khalil jt, 2016; Teare ja Taks, 2020) on selle väljakutse vastu võtnud ning teinud oma ettepanekuid metodoloogia täiendamiseks. Cooper jt (2019) kirjutavad, et metodoloogia aruteludes on esile tõusnud n-ö Kanada uurijate grupp (Levac jt, 2010; Doudt jt, 2013; Colquhoun jt, 2014; O'Brien jt, 2016; Tricco jt, 2016) ja Austraalia uurijate grupp (Peters jt, 2015, 2017, 2020; Khalil jt, 2016). Saab öelda, et prominentsemad ettepanekud metodoloogia osas on teinud Levac jt (2010) ning Peters jt (2015, 2017, 2020). Levac jt (2010) seadsid oma töö eesmärgiks Arksey ja O'Malley poolt väljapakutud metodoloogilist lähenemist detailsemalt kirjeldada, st nad annavad oma kogemusele tuginedes konkreetseid juhiseid iga kirjandusülevaate etapi läbiviimise osas. Peters jt (2015, 2017, 2020) võtsid oma tähelepanekute ning ettepanekute tegemisel aluseks nii Arksey ja O'Malley raamistiku kui Levac jt ettepanekud ja täpsustavad omalt poolt veelgi erinevate sissejuhatava kirjandusülevaate etappide läbiviimist.

Kuigi Arksey ja O'Malley (2005) seda otseselt välja ei too, siis Levac jt (2010) ning Peters jt (2015, 2017, 2020) soovivad sissejuhatavat kirjandusülevaadet teha uurijate meeskonnas, sest selline lähenemine vähendab tõenäosust, et uurimuse erinevates etappides (nt asjakohase kirjanduse otsimisel, allikate valikukriteeriumite sõnastamisel, andmete analüüsimisel jne) ollakse otsuste tegemisel kallutatud. Käesoleva uurimustöö kontekstis ei ole selline lähenemine

võimalik, kuid olen kõik uurimuse etapid ning nende raames tehtud valikud ja otsused uurimustöö juhendajaga läbi rääkinud. Selleks et tagada uurimisprotsessi läbipaistvus ja korratavus olen iga uurimuse etapi juurde lisanud selgitused erinevate tegevuste teostamise kohta.

2.1.1.1 Uurimusküsimuse sõnastamine

Arksey ja O'Malley (2005) kirjutavad, et sissejuhatav kirjandusülevaade algab uurimusküsimuse sõnastamisega. Nad lisavad, et kuna uurimusküsimus mõjutab uurija valikuid ning otsuseid kirjandusanalüüsi järgmistes etappides, siis on oluline teadvustada, mis on küsimuse juures kõige olulisem, st millel on rõhuasetus ja kuidas see mõjutab uurimuse läbiviimist. Sealjuures soovitavad nad pigem n-ö avarat vaadet, et kirjanduse otsimise etapis oleks võimalik leida võimalikult palju asjakohaseid allikaid. Nagu nad välja toovad, siis järgmistes etappides saab uurija nagunii erinevaid piiranguid seada, seda nii kirjanduse otsimisel kui allikate valikul.

Levac jt (2010) leiavad, et uurimusküsimus ise võib olla n-ö laialt sõnastatud, kuid uurimuse ulatus peaks olema selgelt määratletud. Nende hinnangul lihtsustab see uurija tööd järgmistes etappides. Lisaks sellele soovitavad nad uurijal uurimusküsimuse sõnastamisel sissejuhatava kirjandusülevaate läbiviimise põhjuse peale mõelda, st nende kahe vahel peab olema kooskõla – ka see toetab uurija otsuseid (kirjanduse otsimist, allikate valikut, materjalist kokkuvõtte tegemist) järgmistes etappides.

Nagu ülal väljatoodud, siis olen käesoleva magistritöö läbiviimiseks sõnastanud kaks uurimusküsimust:

- 1) Millised on suurandmete kasutamise võimalused sotsiaaltöös teadusartiklite põhjal?
- 2) Millised on suurandmete kasutamisega seotud väljakutsed sotsiaaltöös teadusartiklite põhjal?

Küsimused olen sõnastanud selliselt, et need oleksid piisavalt avarad nagu Arksey ja O'Malley (2005) soovitavad. Sotsiaaltöö hõlmab endas erinevaid alavaldkondi, nt Eestis on võimalik sotsiaaltöös spetsialiseeruda muuhulgas laste ja peredega, psüühilise erivajadustega inimestega, eakatega, sõltlastega, vähemustega jne töötamisele (Eesti Sotsiaaltöö Assotsiatsioon, i.a). Samas on uurimuse ulatus piisavalt konkreetselt määratletud nagu Levac jt (2010) soovitavad, sest uurin suurandmete kasutamise võimalusi ja suurandmete kasutamisega seotud väljakutseid ühes konkreetses valdkonnas.

2.1.1.2 Asjakohase kirjanduse otsimine

Sissejuhatava kirjandusülevaate teises etapis hakatakse uurimustöö läbiviimiseks kirjandust otsima. Seda tehes peab olema põhjalik, et tuvastada võimalikult palju asjakohaseid allikaid (Arksey ja O'Malley, 2005; Peters jt, 2015, 2017). Samas, Levac jt (2010) lisavad, et uurija peab leidma tasakaalu kirjandusülevaate ulatuse ja põhjalikkuse vahel. Kui allikaid on liiga palju, siis võib tekkida olukord, kus uurijal ei ole võimalik neid põhjalikult analüüsida. Uurimus peab olema teostatav ning allikate otsimisel peab arvestama uurimuse läbiviimisega seotud piirangutega (Levac jt, 2010, Peters jt, 2017). Allikadena võib kasutada nii teadusartikleid kui n-ö halli kirjandust (nt avaldatud ja avaldamata raportid jms) ning allikad võivad olla väga erineva metodoloogilise lähenemisega, mille kvaliteeti sissejuhatava kirjandusülevaate läbiviimise raames ei hinnata (Arksey ja O'Malley, 2005; Levac jt, 2010; Colquhoun, 2014; Tricco, 2016; Peters, 2015).

Kirjanduse otsimisel on oluline läbi mõelda ja paika panna allikate otsimise strateegia (Arksey ja O'Malley, 2005; Levac jt, 2010; Peters jt, 2015, 2017), mis on kooskõlas uurimuse eesmärgi ja uurimusküsimusega (Levac jt, 2010). Enne otsingu tegemist on vaja otsustada, milliseid andmebaase selleks kasutatakse ja milliste märksõnadega otsing tehakse (Arksey ja O'Malley, 2005). Siin on uurijale abiks prooviotsingute tegemine, mida soovivad nii Arksey ja O'Malley (2005) kui Peters jt (2015, 2017), sest prooviotsingutega leitud allikate pealkirjade, sisukokkuvõtete ja võtmesõnade lugemisel võib tuvastada täiendavaid märksõnu, mida otsingu tegemisel kasutada. Peters jt (2017) kirjutavad, et sissejuhatava kirjandusülevaate läbiviimisel on kirjanduse otsimine n-ö jooksev tegevus, kus uurija võib uurimuse jooksul allikate otsimise strateegiat täiendada. Lisaks eelnevale on vaja veel otsustada, kas allikatele on vaja rakendada erinevaid piiravaid tingimusi, nt seada piirangud allikate avaldamise ajale, keelele vms (Arksey ja O'Malley, 2005; Peters jt, 2015, 2017).

Kokkuvõtvalt saab öelda, et kõik uurijad peavad erinevatel praktilistel põhjustel (nt uurimusmeeskonna suurus, uurimuse läbiviimiseks olemasolev aeg jms) kirjanduse otsimise etapis piiranguid seadma. Sealjuures on oluline tehtud otsuseid põhjendada ja teadvustada, kuidas seatud piirangud võivad uurimust mõjutada (Levac jt, 2010; Peters jt, 2015, 2017). Mis iganes strateegiat uurija ka ei vali, siis on oluline, et valitud lähenemisega oleks võimalik uurimuse eesmärk saavutada ning uurimusküsimusele vastata (Levac jt, 2010).

Allikate otsimise strateegia määratlemiseks konsulteerisin Tartu Ülikooli raamatukogu erialainfo spetsialistiga, kes andis soovitusi nii andmebaaside kasutamise kui päringu sõnastamise osas. Lisaks sellele lähtusin otsingustrateegiat määrates Tartu Ülikooli raamatukogu (i.a) andmebaaside kasutamise juhendmaterjalist. Täiendavat tuge pakkus ka uurimustöö juhendaja, kellega arutasime muuhulgas, milliseid otsingusõnu võiks päringutes kasutada.

Asjakohaste allikate leidmiseks kasutasin *Ebsco Discovery* otsingumootorit, sest see võimaldab allikaid otsida korraga erinevates teadusandmebaasides (Tartu Ülikooli raamatukogu..., i.a). Selleks, et oma otsingut paremini suunata kasutasin liitotsingut (ingl *advanced search*), mis on täpsem kui otsingumootori poolt vaikimisi ette antud lihtotsing (ingl *basic search*) (Tartu Ülikooli raamatukogu..., i.a). Vastavalt TÜ raamatukogu juhendmaterjalile saab liitotsinguga muuhulgas täpsustada kriteeriume, millele allikad vastama peavad, nt on võimalik allikaid otsida allika ilmumisaaja, keele, väljaande tüübi jt kriteeriumite järgi. Samuti saab liitotsinguga täpsustada, kas otsingusõna peaks esinema allika pealkirjas, sisukokkuvõttes, täistekstis, märksõnade hulgas jne.

Päringu sõnastamisel kasutasin Boole'i loogikaoperaatoreid OR ja AND, kus OR võimaldab otsingut laiendada ning AND otsingut piirata. Lisaks sellele panin kõik otsingusõnad ja otsingusõnade kombinatsioonid jutumärkidesse, et leida allikaid, mis sisaldaksid just selliseid sõnu ja sõnade kombinatsioone nagu olen allpool välja toonud. Täiendavalt kasutasin sümbolit *, et oleks võimalik leida erinevate lõppudega sõnu (nt *social work** - leitakse allikad, mis sisaldavad nii sõnu *social work* (sotsiaaltöö) kui *social worker* (sotsiaaltöötaja).

Prooviotsinguid teostasın jaanuaris ja veebruaris 2021 ning esimese päringu tegin 26.02.2021 järgmiste otsingusõnadega: sotsiaalpoliitika (ingl *social policy*), sotsiaaltöö (ingl *social work**), sotsiaalkaitse (ingl *social protection*), sotsiaalhoolekanne (ingl *social welfare*), sotsiaalkindlustus (ingl *social insurance, social security*), sotsiaalteenus (ingl *social service**), lastekaitse (ingl *child welfare, child* protecti**) ja suurandmed (ingl *big data*). Uurimuse eesmärgist ja uurimusküsimustest tulenevalt lähtusin esimese päringu tegemisel otsingusõnade valikul põhimõttest, et päringuga oleks võimalik leida allikaid, mis käsitleksid suurandmete kasutamist sotsiaaltöös kogu valdkonna mitmekesisuses. Esimese päringu tulemusel sain 32 allikat, millest 17 olid unikaalsed allikad.

Pärast esimese päringuga saadud ning pealkirja ja sisukokkuvõtte põhjal täisteksti lugemise etappi liikunud allikate läbivaatamist lisasin otsingusõnade hulka masinõppe (ingl *machine*

learning), andmekaeve (ingl *data mining*) ja algoritmi (ingl *algorithm**). Tegemist on märksõnadega, mis olid välja toodud kas erinevate allikate märksõnade hulgas või täistekstides endis. Teise päringu koos täiendavate otsingusõnadega tegin 26.03.2021. Päringu tulemusel sain 92 allikat, millest 39 olid unikaalsed allikad.

Lisaks sellele otsustasin teha kolmanda päringu, kus lisasin otsingusõnade hulka veel infosüsteemi(d) (ingl *information system**), elektroonilise(d) infosüsteemi(d) (ingl *electronic information system**), otsuste tegemise (ingl *decision making*), automatiseeritud otsuste tegemise (ingl *automated decision making*), otsuste tegemist toetav(ad) abisüsteem(id) (ingl *decision support tool*, decision support system*, algorithmic tool*, predictive tool**) ning andmebaasi(d) (ingl *database**). Sarnaselt teisele päringule lähtusin kolmandat päringut tehes otsingusõnade ja otsingusõnade kombinatsioonide valikul esimese ja teise päringuga leitud allikate pealkirjade ja sisukokkuvõtete põhjal täisteksti lugemise etappi liikunud allikates kasutatud märksõnadest. Kolmanda päringu koos täiendavate märksõnadega teostasin 06.04.2021. Päringu tulemusel sain 12 allikat, millest ükski ei olnud enam unikaalne allikas.

Kokkuvõtvalt, kolme päringu tulemusel sain kokku 136 allikat, millest 56 allikat olid unikaalsed artiklid, mida hakkasin uurimustöö järgmises etapis täiendavalt läbi vaatama. Päringud koos otsingusõnade ja otsingusõnade kombinatsioonidega ning tulemustega on esitatud tabelis 1.

Tabel 1. Päringud koos otsingusõnade ja otsingusõnade kombinatsioonide ning tulemustega

Andmebaas Ebsco Discovery	Otsing	Otsingusõnad ja otsingusõnade kombinatsioonid	Tulemus
	I otsing	"social policy" OR "social work*" OR "social protection" OR "social welfare" OR "social insurance" OR "social security" OR "social service*" OR "child welfare" OR "child* protecti*" AND „big data“	32
	II otsing	"social policy" OR "social work*" OR "social protection" OR "social welfare" OR "social insurance" OR "social security" OR "social service*" OR "child welfare" OR "child* protecti*" AND “big data” OR “data mining” OR “machine learning” OR “algorithm*”	92
	III otsing	"social policy" OR "social work*" OR "social protection" OR "social welfare" OR "social insurance"	12

		OR "social security" OR "social service*" OR "child welfare" OR "child* protecti*" AND "big data" OR "data mining" OR "machine learning" OR "algorithm*" AND "decision making" OR "automated decision*" OR "automated decision making" OR "decision support system*" OR "decision support tool*" OR "algorithmic tool*" OR "predictive tool*" OR "information system*" OR "electronic information system*" OR "database*"	
Kokku			136

2.1.1.3 Allikate valimine

Allikate valikul tuginetakse kaasamis- ja välistavatele kriteeriumitele, mille sõnastamisel on lähtutud uurimusküsimusest ja mida võib uurimuse käigus vastavalt vajadusele muuta (Arksey ja O'Malley, 2005; Levac jt, 2010; Peters jt, 2017). Seega saab öelda, et ka allikate valimine on n-ö jooksev tegevus, kus kriteeriumite muutmise vajadus võib selguda pärast seda, kui uurija on jõudnud kirjandusega tutvuda.

Täiendavalt, ka siin etapis võib olla vaja täiendavaid piiranguid seada. Näiteks Arksey ja O'Malley (2005) kirjutavad, et nemad pidid selle etapi läbimiseks seadma ajalise piirangu pärast mida nad rohkem allikaid ei analüüsinud. Küll aga lisasid nad info kõikide mitte analüüsitud allikate kohta oma uurimusele, et järgmistel uurijatel oleks vajaduse korral võimalik nendega tutvuda.

Sarnaselt kirjanduse otsimise etapile peab allikate valiku protsess olema läbipaistev ja korratav, mistõttu on oluline, et uurija põhjendaks tehtud otsuseid (Levac jt, 2010; Peters jt, 2015, 2017). Peters jt (2015) soovivad uurimusele lisada allikate otsinguprotsessi kirjeldava diagrammi (nt PRISMA diagramm), kus oleks koos selgitustega välja toodud leitud allikate arv ning lõpliku valimi moodustunud allikate arv koos vahepealsete etappidega, kus on välja toodud allikate valiku põhimõtted ja kust on näha, kuidas on lõpliku valimini jõutud.

Kui tegin prooviotsinguid, siis ma ei seadnud päringule esialgu piiravaid tingimusi. Selle tulemusel sain tulemuseks väga suure arvu allikaid, mida vaadates oli selge, et nende hulgas

on minu uurimustöö kontekstis palju mitte relevantset kirjandust. Arvestades käesoleva magistritöö läbiviimiseks olemasolevat aega ei oleks mul olnud võimalik kõiki päringuga leitud allikaid läbi vaadata ning analüüsida. Tartu Ülikooli raamatukogu andmebaaside juhendmaterjalis (i.a) kirjutatakse, et kui päringu tulemusel leitakse liiga palju tulemusi, siis on eesmärgipärane päringut piirata, nt allikate ilmumisaastate ja pealkirja või sisukokkuvõtte välja järgi. Kuna kõik allikate sisu kirjeldavad olulised sõnad peaksid esinema kas sisukokkuvõttes või pealkirjas, siis on päringut selliselt piirates võimalik jõuda kõige asjakohasema kirjanduseni (Tartu Ülikooli raamatukogu..., i.a).

Sellest tulenevalt otsustasin päringu teha selliselt, et kõik mulle huvipakkuvad otsingusõnad esineksid allikate pealkirjas. Lisaks sellele piirasin päringut allikate ilmumisaastatega. Arvestades kui kiiresti tehnoloogia, sh suurandmetega seonduv (nt erinevad võimalused suurandmeid koguda, kasutada ja analüüsida) areneb, siis otsisin allikaid, mis on avaldatud perioodil 2015 kuni 06.04.2021 ehk kuni kolmanda päringu tegemise kuupäevani, et leida võimalikult uut kirjandust. Allikate kaasamis- ja välistavad kriteeriumid olen välja toonud tabelis 2.

Tabel 2. Kaasamise- ja välistavad kriteeriumid

Kaasamise kriteeriumid	Välistavad kriteeriumid
Otsingusõnad esinevad allika pealkirjas	Otsingusõnad ei esine allika pealkirjas.
Allika täistekst on kättesaadav	Allika täistekst ei ole kättesaadav
Allikas on ilmunud ajavahemikus 2015–06.04.2021, st kuni kolmanda päringu tegemise kuupäevani	Allikas on ilmunud enne 2015. aastat või pärast 06.04.2021, st pärast kolmanda päringu tegemist
Allikas on kirjutatud inglise keeles.	Allikas ei ole kirjutatud inglise keeles.
Allikas käsitleb suurandmete kasutamist sotsiaaltöös (nt suurandmeid kasutatakse otsuste tegemiseks, sotsiaalteenuste planeerimiseks vms)	Allikas ei käsitle suurandmete kasutamist sotsiaaltöös.

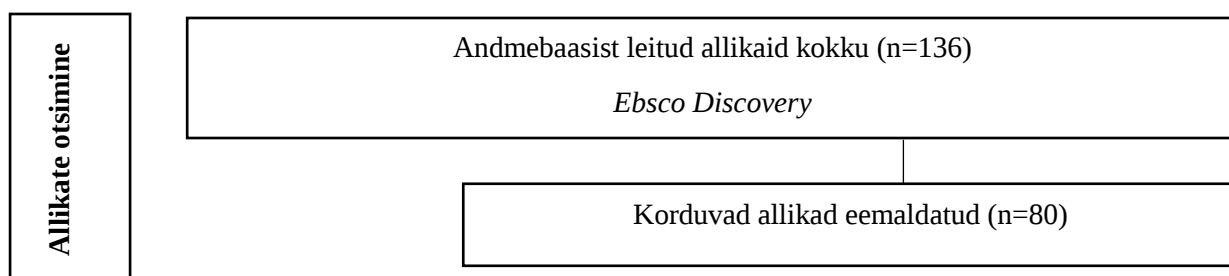
Nagu eelmises alapeatükis välja tõin, siis kolme päringu tulemusel sain kokku 136 allikat. Pärast korduvate allikate eemaldamist jäi järgi 56 unikaalset allikat, mille asjakohasust hakkasin vastavalt kaasamise- ja välistavatele kriteeriumitele allikate pealkirja, sisukokkuvõtte ja täisteksti lugemise põhjal hindama (vt joonis 1).

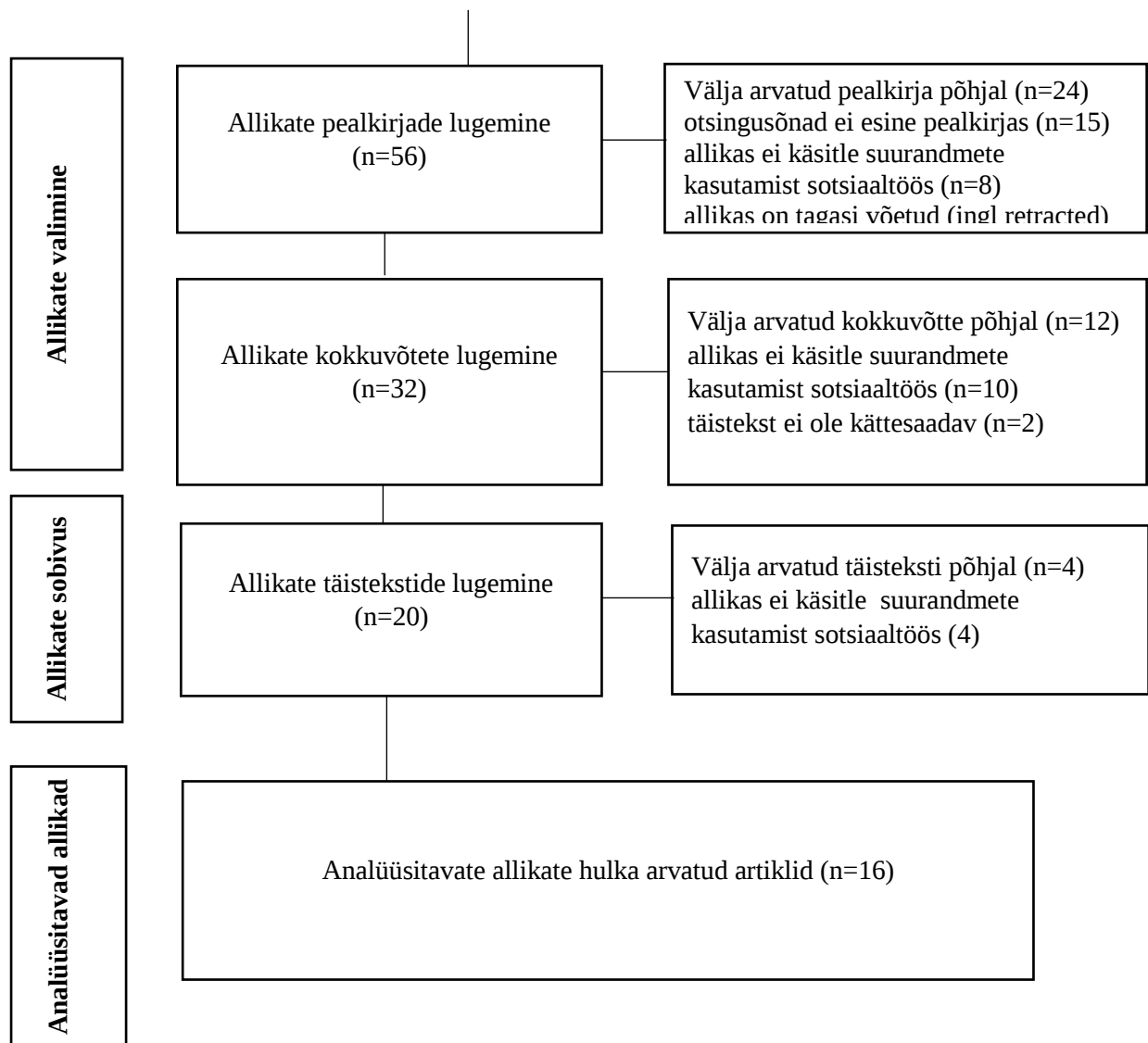
Allikate esmase valiku tegin pealkirja järgi. Esimese ja teise päringu puhul arvasin sisse kõik allikad, mille pealkirjas esinesid uurimustöö otsingusõnad. Kuna kõik kolmanda päringuga leitud allikad juba kajastusid esimese ja teise päringu tulemuste hulgas, siis lähtusin nende asjakohasuse hindamisel eespool öeldust. Lisaks sellele arvasin sisse kõik allikad, mille puhul oli pealkirja põhjal võimalik öelda, et need käsitlevad suurandmete kasutamist sotsiaaltöös või vastupidi, mille puhul ei olnud pealkirja põhjal võimalik öelda, kas need käsitlevad suurandmete kasutamist sotsiaaltöös. Siit edasi, pealkirjade lugemise põhjal jätsin välja sellised allikad, mille puhul oli pealkirja põhjal võimalik öelda, et need ei käsitle suurandmete kasutamist sotsiaaltöös.

Seejärel lugesin läbi pealkirja põhjal sisse arvatud allikate kokkuvõtted. Samaselt ülal kirjeldatule arvasin sisse kõik allikad, mille puhul oli kokkuvõtte põhjal võimalik öelda, et allikas käsitleb suurandmete kasutamist sotsiaaltöös ning vastupidi, st allikad, mille puhul kokkuvõtte põhjal ei olnud võimalik aru saada, kas allikas käsitleb soovitud teemat või mitte. Samuti arvasin sisse allikad, millel kokkuvõtte puudus, st allika asjakohasust oli võimalik hinnata üksnes täisteksti lugemisega. Välja jätsin omakorda allikad, mille puhul oli kokkuvõtte põhjal võimalik öelda, et need ei käsitle suurandmete kasutamist sotsiaaltöös.

Lõpliku valimi moodustamiseks lugesin läbi pealkirja ja sisukokkuvõtte lugemise etapist täisteksti lugemise etappi liikunud allikad. Täistekstide põhjal arvasin sisse allikad, mis olid kooskõlas uurimustöö eesmärgi ja küsimustega. Välja jätsin sellised allikad, mille lugemisel selgus, et need siiski ei aita saavutada uurimustöö eesmärki ning vastata uurimustöö küsimustele. Samuti jäid välja allikad, mille täistekst ei olnud kättesaadav.

Kokkuvõtvalt, allikate esmase valiku nii pealkirja, kokkuvõtte kui täisteksti põhjal tegin ise. Selleks, et vältida vigu ja tagada tehtud valikute asjakohasus vaatas uurimustöö juhendaja minu poolt tehtud valikud koos selgituste ja põhjendustega üle. Lõplikud otsused rääkisime läbi. Kuigi ma ei piiranud oma otsingut allika liigi järgi (ingl *source type*), siis moodustavad lõpliku valimi 16 teadusartiklit (kasutan edaspidi sünonüümina ka kirjandus, allikas, artikkel). Allikate valimisprotsessi visualiseerimiseks kasutasin Prisma diagrammi (joonis 1).





Joonis 1: PRISMA diagramm: allikate valimisprotsess (Autori joonis PRISMA *transparent reporting of systematic reviews and meta-analyses* 2021 järgi)

2.1.1.4 Andmete kaardistamine

Neljas oluline etapp sissejuhatava kirjandusanalüüsi läbiviimisel on andmete kaardistamine enne mida on vaja otsustada, millist infot hakatakse valimi moodustanud allikatest koguma. Selleks loob uurija n-ö andmete kogumise vormi, kuhu pannakse kirja allikate põhinäitajad, milleks võib muuhulgas olla info uurimuse autori, uurimuse avaldamise aasta, uurimuse eesmärgi, metodoloogia, uurimuse sihtrühma ja sekkumise, uurimuse tulemuste jms kohta (Arksey ja O'Malley, 2005; Levac jt, 2010; Peters jt, 2015, 2017).

Levac jt (2010) ning Peters jt (2015, 2017) lisavad, et koguma peab selliseid andmeid, mis võimaldavad uurimusküsimusele vastata ja uurimuse eesmärgi täita. Seega saab öelda, et andmete kogumise vormid võivad olla väga erinevad, sest uurija peab sinna lisama selliseid andmevälju, mis on konkreetse uurimustöö kontekstis asjakohased ja vajalikud. Levac jt (2010) ning Peters jt (2015) soovivad loodud vormi esmalt mõne uurimuse peal testida, et näha, kas seda on vaja muuta või mitte, kuid üldiselt leiavad mõlemad autorid, et andmete kogumise vormi võib kogu uurimuse vältel vastavalt vajadusele täiendada.

Käesoleva uurimustöö andmete kogumise vorm (vt lisa 1) sisaldab järgmisi andmevälju:

- Teadusartikli pealkiri
- Ajakiri, kus teadusartikkel on avaldatud
- Autor
- Teadusartikli avaldamise aasta
- Riik, kus uurimus läbi viidi või kus teadusartikkel avaldati
- Andmete liik, st milliseid andmeid uurimuses kasutatakse või milliste andmete võimalikku kasutamist käsitletakse
- Sotsiaaltöö valdkond, kus suurandmeid kasutati või kus neid saaks kasutada
- Kas teadusartikkel käsitleb suurandmete kasutamisega seotud eetilisi küsimusi
- Kas teadusartikkel sisaldab praktilisi näiteid suurandmete kasutamise kohta

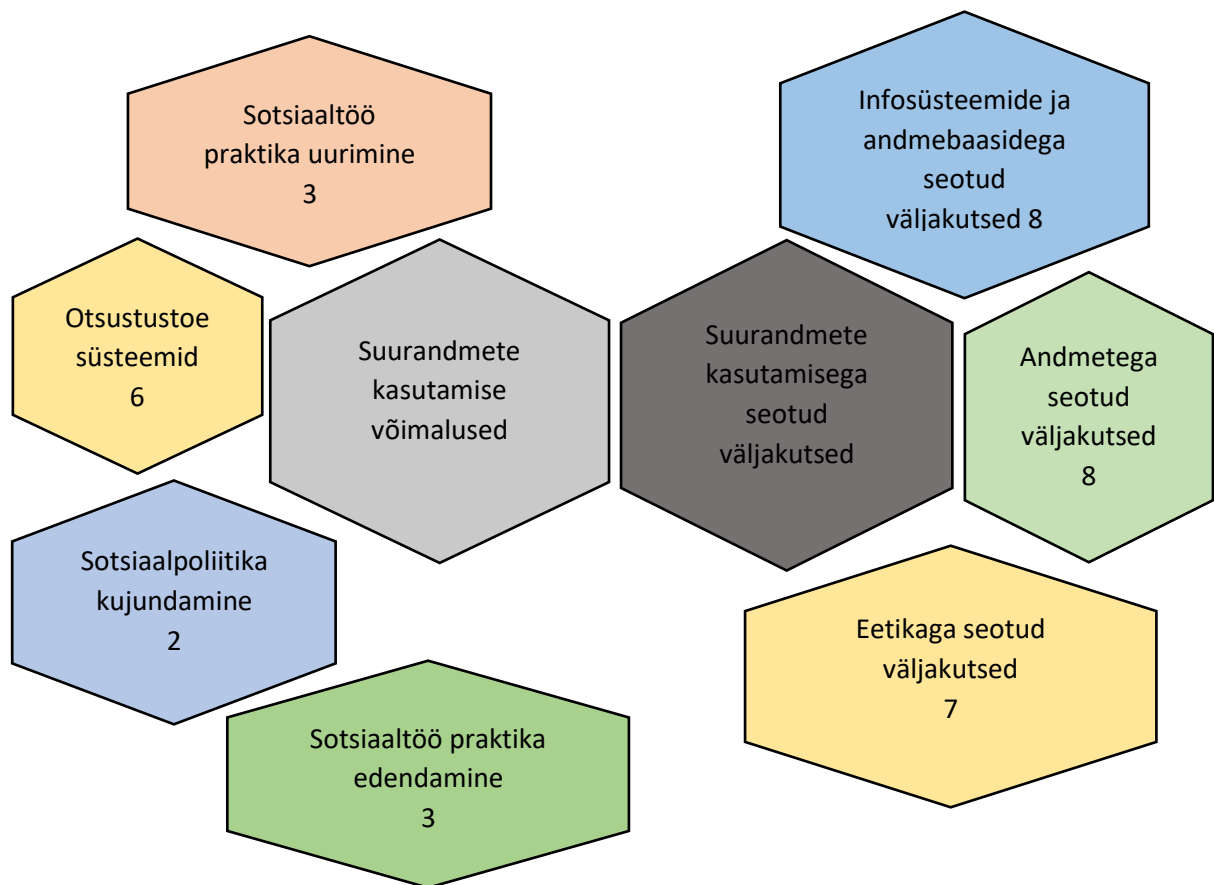
2.1.1.5 Materjali läbitöötamine, kokkuvõtte tegemine, tulemuste esitlemine

Sarnaselt andmete kaardistamise etapile (vt alapeatükk 2.1.1.4) on ka siin etapis oluline läbi mõelda, kuidas läbitöötatud materjalist kokkuvõtte teha ning millisel moel tulemusi esitleda (Arksey ja O'Malley, 2005). Siinjuures on abiks eelmises etapis loodud andmete kogumise vorm, mis võimaldab analüüsitud allikatest anda nii numbrilise kui temaatilise ülevaate (Arksey ja O'Malley, 2005). Samas, Peters jt (2015) lisavad, et see, kuidas uurimustulemusi esitleda on lõpuks iga uurija enda otsustada sõltudes muuhulgas uurimusküsimusest ja uurimuse eesmärgist. Sellegipoolest annavad nii Peters jt (2015, 2017) kui Levac jt (2010) siin uurijatele täiendavaid soovitusi.

Näiteks, Levac jt (2010) soovivad uurijatel keskenduda kolmele tegevusele: andmete analüüsimisele, tulemuste esitlemisele ning tulemuste lahti mõtestamisele. Esiteks, nad leiavad oma kogemusele tuginedes, et allikate analüüsimisel on sobiv kasutada kvalitatiivsed andmeanalüüsi meetodeid (nt kvalitatiivset sisuanalüüsi). Teiseks, nad toovad välja, et tulemuste esitlemisel on oluline lähtuda sissejuhatava kirjandusülevaate eesmärgist ning valida sellest tulenevalt sobivaim viis tulemuste kajastamiseks. Ka Peters jt (2015) rõhutavad tulemuste esitlemise ja sissejuhatava kirjandusülevaate eesmärgi kooskõla olulisust lisades, et tulemuste kajastamisel võib muuhulgas kasutada erinevaid formaate. Kolmandaks, Levac jt (2010) peavad oluliseks, et uurimustulemused asetataks laiemasse konteksti ja analüüsitakse, milline on uurimustulemuste tähendus/mõju uurimusele, poliitikale ja praktikale.

Uurimustöö lõplik valim koosneb 16 teadusartiklist. Enamik uurimustest on läbi viidud või avaldatud Ameerika Ühendriikides (kokku 7 artiklit). Järgnesid Austraalia (kokku 4 artiklit) ning Uus-Meremaa, Rootsi, Saksamaa, Hiina ning Korea kõik ühe artikliga. Kõikides valimi moodustanud teadusartiklites käsitleti sotsiaaltöö valdkonna suurandmetena administratiivseid andmeid. Koreas läbiviidud uurimus (Choo, Joo, Kim ja Myung, 2019) tõi lisaks administratiivsete andmete kasutamise võimalusele välja sensorite kogutavate andmete, ruumiandmete, mobiilirakenduste ja veebirakenduste ning sotsiaalmeedia andmete (nt Twitter ja Kakao Talk, viimane on Koreas kasutusel olev sõnumirakendus) aga ka telefonilogi andmete kasutamise võimaluse ehkki uurimuses endas kasutati üksnes administratiivseid andmeid. 10 teadusartiklis tuuakse näiteid sellest, kuidas on suurandmeid sotsiaaltöö valdkonnas kasutatud või kuidas oleks suurandmeid võimalik kasutada; 5 artiklit on empiirilised uurimused, kus uurimuse läbiviimisel kasutati suurandmeid ning üks artikkel on teoreetiline artikkel, kus räägitakse andmete kasutamise põhimõtetest ilma suurandmete kasutamise näiteid välja toomata.

Suurandmete kasutamisega seotud võimalusi sotsiaaltöös võib käsitleda nelja kategooria lõikes, milleks on: otsustustoe süsteemid (ingl *decision support systems, decision support tools, algorithmic tools, predictive tools*), sotsiaaltöö praktika uurimine, sotsiaaltöö praktika edendamine ning sotsiaalpoliitika kujundamine. Suurandmete kasutamisega seotud väljakutseid võib seevastu käsitleda kolme kategooria lõikes, milleks on: infosüsteemidest ja andmebaasidest tulenevad väljakutsed; andmetega (sh andmete kvaliteediga) seotud väljakutsed ning eetiliste küsimustega seotud väljakutsed.



Joonis 2: tulemuste kaardistus kategooriate ja allikate arvu kaupa (Joonis Peters jt 2015 joonise järgi).

2.1.1.6 Konsultatsioon

Kuues ja viimane etapp *scoping review* läbiviimisel on konsultatsioon huvirühmade seas. Arksey ja O'Malley (2005) sissejuhatava kirjandusülevaate metodoloogilise lähenemise kirjelduses on tegemist valikulise etapiga. Käesolevas uurimustöös konsultatsiooni etappi läbi ei viidud.

3 ANALÜÜS

3.1 Suurandmete kasutamise võimalused sotsiaaltöös kirjanduse põhjal

Käesolevas peatükis kirjeldan, millised on suurandmete kasutamise võimalused ja suurandmete kasutamisega seotud väljakutsed sotsiaaltöös uurimustöö valimi moodustanud teadusartiklite põhjal. Analüüsitud artiklites käsitletakse sotsiaaltöö valdkonna suurandmetena infosüsteemides ja andmebaasides talletatud andmeid.

3.1.1 Otsustustoe süsteemid

Uurimustöö valimi moodustanud allikate hulgas on kõige rohkem selliseid artikleid, mis räägivad suurandmete kasutamisest otsustoe süsteemide (kasutan sünonüümina ka otsuste tegemise abisüsteem, abisüsteem, vahend) (ingl *decision support systems, decision support tool, algorithmic tool, predictive tool*) loomisel sotsiaaltöös (Schwartz, York, Nowakowski-Sims ja Ramos-Hernandez, 2017; Keddell, 2019; Gillingham, 2019a, 2019b; Schneider ja Seelmeyer, 2019; Gillingham, 2020). Kahes allikas (Lery, Haight ja Alpert, 2016; Choo, Joo, Kim, Cha ja Myung, 2018) ei kasutata terminit otsustustoe süsteem, kuid Lery jt (2016) kirjutavad suurandmete kasutamise põhimõtetest, et toetada otsuste tegemist lastekaitse-süsteemis ning Choo jt (2018) kirjeldavad, kuidas nad kasutasid suurandmeid, et välja töötada laste väärkohtlemist ennetav süsteem. Sellest tulenevalt olen mõlemad allikad välja toonud käesoleva kategooria all. Tervikuna, analüüsitud kirjanduses käsitletakse mitmeid erinevaid teemasid seoses suurandmete kasutamise ja analüüsimisega otsuste tegemise toetamise eesmärgil. Need on seotud otsustustoe süsteemide toimimise aluseks olevate andmetega, sh andmete kvaliteedi ja andmekaitset puudutavate küsimustega (Gillingham, 2020; Keddell, 2019; Schneider ja Seelmeyer, 2019), infosüsteemidega (Gillingham, 2019), otsustustoe süsteemide täpsuse, kasulikkuse ja läbipaistvusega, sotsiaaltöö põhimõtete, väärtuste ning eetikaga jpm (Gillingham, 2019b; Keddell, 2019; Schneider ja Seelmeyer, 2019).

Otsustustoe süsteemid on n-ö uue põlvkonna tööriistad, mis erinevad sotsiaaltöös kasutatavatest traditsioonilistest lähenemistest (nt paberil või elektroonselt läbiviidavatest riskihindamistest, mis samuti toetavad otsuste tegemist) selle poolest, et need tuginevad andmetehnoloogiate võimalustele (Gillingham, 2019b). Siinjuures on oluline, et abisüsteemi eesmärk ei ole otsuseid sotsiaaltöötaja eest ära teha, vaid toetada sotsiaaltöötajat otsuste

tegemisel (Gillingham, 2019b; Scheider ja Seelmeyer, 2019). See tähendab, tegemist ei ole automatiseeritud otsuste tegemisega, kus otsus tehakse ilma inimese sekkumiseta, vaid info kokku koondamisega selliselt, et sotsiaaltöötaja saaks seda otsuste tegemisel kasutada (Scheider ja Seelmeyer, 2019).

Sotsiaaltöö valdkonnas on hakatud üha enam selliseid andmetehnoloogiatele tuginevaid otsustustoe süsteeme välja töötama ja rakendama ning sellega seoses kirjeldatakse käesolevas magistritöös analüüsitud teadusartiklites mitmeid murekohti, mis on seotud nii abisüsteemide loomise kui abisüsteemi poolt tehtud soovitusetega ja nende rakendamisega praktikas. Gillingham (2020), tuginedes *The United States Association for Computing Machinery* (2017) poolt tehtud pöördumisele algoritmide läbipaistvuse ja vastutuse teemal (ingl *Statement on Algorithmic Transparency and Accountability*) selgitab, et pöördumises väljatoodud põhimõtted on asjakohased ka sotsiaaltöö valdkonnas ning võiksid seega pakkuda n-ö toetava raamistiku nii otsustustoe süsteeme kasutavatele asutustele kui sotsiaaltöötajatele. Nimetatud põhimõtted on sõnastatud järgmiselt: 1) teadlikkus, st algoritmide loojad ja kasutajad peavad olema teadlikud võimalikest algoritmide eelarvamustest ja nende tagajärgedest inimestele, gruppidele ja ühiskonnale, 2) juurdepääs ja vea parandamise võimalus, st asutustes, mis kasutavad algoritme peavad olema loodud võimalused algoritmide poolt tehtud otsuste vaidlustamiseks, 3) vastutus, st inimene, kes teeb algoritmi soovitusel tuginedes lõpliku otsuse peaks vastutama tehtud otsuse eest, 4) selgitus, st asutused ja spetsialistid, kes kasutavad otsuste tegemisel algoritme peavad olema võimelised selgitama, kuidas algoritm konkreetsete soovituseni jõudis, 5) andmete päritolu, st peab olema selge ülevaade, kuidas on andmed, millega algoritmi on õpetatud, kogutud ja millised on võimalikud andmetes sisalduvad eelarvamused, 6) auditeeritavus, st algoritmide loojad ja kasutajad peavad alles hoidma dokumendid, mis näitavad, kuidas erinevad mudelid on loodud, milliseid andmeid selleks on kasutatud ja milliseid otsuseid on tehtud, 7) valideerimine ja testimine, st loodud mudeleid peab pidevalt valideerima ja testima, et oleks selge, kas mudelite kasutamine viib diskrimineerimiseni või kahju tekitamiseni, sh peab sellega seotud andmeid alles hoidma.

Andmetehnoloogiatele tuginevate otsustustoe süsteemidega seoses on oluline välja tuua, et sellist süsteemi saab kasutada nii üks-ühele klienditöös kui ka laiemalt sekkumiste planeerimiseks. Kuna abisüsteemi kasutamisel ei ole inimene see, kes andmeid tõlgendab (Gillingham, 2020), siis nähakse selliste otsuste tegemist toetavate süsteemide kasutamise võimalike kasuteguritena muuhulgas seda, et need on efektiivsed (Scheider ja Seelmeyer, 2019), sh võimelised analüüsima rohkem andmeid, seda kiiremini ja täpsemalt kui inimesed,

mistõttu võiks see kasulik olla näiteks sotsiaaltöötaja ja kliendi esmakohatumisel, kus sotsiaaltöötajal on otsuste tegemiseks tihti vähe informatsiooni ja vähe aega (Keddell, 2019). Samuti saab andmetehnoloogiatele tuginevaid abisüsteeme kasutada erinevate teenuste (Schneider ja Seelmeyer, 2019) ja ennetavate tegevuste planeerimiseks (Gillingham, 2020; Keddell, 2019).

Andmetehnoloogiad on erinevad (sh erineva läbipaistvusega) ning andmed, mille põhjal otsustustoe süsteem soovitusi teeb pärinevad erinevatest allikatest (sh on andmete kvaliteet erinev) (Keddell, 2019). Lisaks sellele on erinevatel riikidel erinevad võimalused ja erinev võimekus andmeid kasutada ja omavahel kombineerida (Keddell, 2019). Seda ilmestavad kirjanduses käsitletud praktilised näited otsustustoe süsteemidest, mida on kasutatud või mida kasutatakse nii Ameerika Ühendriikides, Austraalias, Uus-Meremaal, Saksamaal, Hollandis kui Inglismaal (Schwartz, York, Nowakowski-Sims ja Ramos-Hernandez, 2017; Gillingham 2019a, 2019b, 2020; Keddell, 2019; Scheider ja Seelmeyer, 2019). Kirjanduse analüüsi põhjal saan öelda, et otsustustoe süsteemid on leidnud laialdast rakendust just lastekaitse-süsteemis ehkki artiklites on paar praktilist näidet vahenditest, mida kasutatakse teistes sotsiaaltöö valdkondades – ühel juhul selleks, et ennustada perevägivalla toimepanemise eest karistatud meeste retsidiivsust (Gillingham 2019a, 2019b, 2020; Keddell, 2019) ja teisel juhul selleks, et toetada puuetega inimestele teenuste planeerimist (Scheider ja Seelmeyer, 2019). Olenemata otsustustoe süsteemi kasutamise eesmärgist ja rakendusvaldkonnast on otsustustoe süsteemidega seoses mitmeid sotsiaaltöö kontekstis olulisi ning teadvustamist vajavaid aspekte.

Schneider ja Seelmeyer (2019) kirjutavad, et andmetehnoloogiatele tuginevate abisüsteemide puhul leitakse, et need on soovitusi tehes objektiivsed ja erapooletud, st kuna analüüsitakse rohkem ja erinevaid andmeid, siis on tulemused samuti objektiivsemad. Kuna see arusaam on muuhulgas seotud andmetega, mida andmetehnoloogiad kasutavad, siis käsitlen teemat põhjalikumalt allpool suurandmete kasutamisega seotud väljakutsete peatükis, aga olgu siinkohal öeldud, et andmete kvaliteet mõjutab andmete põhjal tehtud soovitusi (Gillingham, 2019b). Siin mängivad oma rolli muuhulgas infosüsteemid, millega seonduvat käsitlen samuti põhjalikumalt väljakutsete peatükis, aga nagu Gillingham (2020) kirjutab, siis on infosüsteemide kasutuselevõtt sotsiaaltöös oluliselt muutnud, mida ja kuidas dokumenteeritakse. Näiteks ta selgitab, et infosüsteemid on tihti üles ehitatud selliselt, et need ei võimalda inimesega seotud asjaolusid kirja panna sellise detailsuse ja täpsusega nagu seda

sai teha pabertoimikutes. Seega annavad kirja pandud andmed üksnes osalise ülevaate inimesest ja tema reaalsest olukorrast (Gillingham, 2020).

Täiendav diskussioon käib otsustustoe süsteemide täpsuse, kasulikkuse ja läbipaistvuse üle. Käesolevas magistritöös analüüsitud kirjanduses väljatoodud näited ilmestavad neid teemasid. Mis puudutab abisüsteemide täpsust, siis Keddell (2019) kirjutab, et võrdlevaid täpsusuuringuid on vähe. Nende uuringute tulemused, millele ta oma töös viitab näitavad, et osade otsustustoe süsteemide puhul on leitud, et need on täpsemad kui näiteks n-ö tavapäraselt kasutatavad riskihindamised või spetsialisti poolt antud hinnangud aga on ka neid uuringuid, kus on leitud, et spetsialistide hinnangud ja otsused on täpsemad kui mistahes abisüsteemi omad. Üks näide sellest, kus leiti, et ennustav ja ettekirjutav analüüs ning andmetehnoloogiad parandavad traditsioonilise lastekaitse süsteemis kasutusel oleva riskihindamise instrumendi täpsust on Schwartz jt (2017) poolt läbi viidud uurimus. Samas on otsustustoe süsteemide täpsusega seonduvat keeruline hinnata ja analüüsida, sest info selle kohta, kuidas sellised vahendid on loodud, milliseid andmeid nad kasutavad ja kuidas täpselt konkreetsete soovituseni jõuavad ei ole tihti avalik (Gillingham, 2019b). Täpsusega seoses kirjutatakse veel, et otsustustoe süsteemid ei arvesta ennustusi ja soovitusi tehes laiemas inimest ümbritseva kontekstiga (Gillingham, 2011; Goddard jt, 1999; Munro jt, 2014, viidatud Keddell 2019). Samas on see sotsiaaltöös oluline, näiteks Lery, Height ja Alpert (2016) kirjutavad, et otsuseid tehes peab arvestama nii keskkonnast tulenevate (nt piirkondlikud näitajad, teenuste kättesaadavus) kui individuaalsete ja juhtumiga seotud teguritega ning samuti selle kontekstiga, kus otsuseid tehakse, sh valdkonna eripärade ja konkreetse organisatsiooniga, kus abisüsteemi kasutatakse (Keddell, 2019).

Mis puudutab andmetehnoloogiatele tuginevate otsustustoe süsteemide kasulikkust, siis Keddell (2019) annab selles osas põgusa ülevaate. Näiteks kirjutab ta, et lastekaitsetöötajate arvamusi käsitlevatest uuringutest on selgunud, et lastekaitsetöötajate hinnangul keskenduvad abisüsteemid liiga palju riskidele ja vähem kaitsvatele teguritele, sh kalduvad need teatud tegurite olemasolu korral riski üle hindama (nt kui peres kasvab suur arv lapsi, kelle vanusevahe on ühtlasi väike jms). Seega otsustustoe süsteemid ei arvesta erinevate tegurite koosmõju individuaalsete juhtumite puhul (Keddell, 2019). Samuti on uuringutest selgunud, et lastekaitsetöötajad ei järgi alati süsteemi poolt tehtud soovitusi. Viimase osas kirjutab Gillingham (2019b), et on vähe kirjandust, mis käsitleks, kuidas sotsiaaltöötajad saavad otsustustoe süsteemi poolt tehtud soovitusi vaidlustada, sh kuna ja millistel põhjustel nad peaksid seda tegema.

Täiendavate oluliste teemadena käsitletakse kirjanduses sotsiaaltöö põhimõtete, väärtuste ja eetikaga seonduvat. Keddell (2019) tuginedes rahvusvahelisele sotsiaaltöö föderatsioonile (2014) kirjutab, et sotsiaaltöös on olulised austus inimeste vastu ja inimõigused ning humanne ja suhetel põhinev lähenemine praktikale. Nimetatud põhimõtete ja andmetehnoloogiatele tuginevate otsustustoe süsteemide vahel on erinevaid vastuolusid (Keddell, 2019; Schneider ja Seelmeyer, 2019; Gillingham, 2019b), mida käsitlen põhjalikumalt suurandmete kasutamisega seotud väljakutsete peatükis.

3.1.2 Sotsiaaltöö praktika uurimine

Allikates, mis käsitlevad suurandmete kasutamist sotsiaaltöö praktika uurimisel kasutati andmekaevet, et administratiivseid andmeid (antud uurimustes sotsiaaltöötajate hinnanguid ja lastekaitsetöötajate juhtumitega seotud dokumente, ingl. k. *case records*) analüüsida ning seeläbi praktilise sotsiaaltöö ja lastekaitsetöö kohta teadmisi saada.

Kaks neist (Carnochan jt, 2018; Taylor jt, 2019) on läbi viidud Ameerika Ühendriikides ning üks (Larsson, Fredriksson ja Sjogren Fugl-Meyer, 2019) Rootsis. Carnochan jt (2018) analüüsisid lastekaitsetöötajate juhtumitega seotud dokumente (sh lastekaitsetöötajate koostatud juhtumikirjeldusi, juhtumiplaane, hinnanguid ja märkmeid ning uurimismenetluse käigus kogutud andmeid ja kohtudokumente), et nende põhjal välja selgitada, milline on n-ö tipptasemel lastekaitsetöö olemus (ingl. k. *nature of excellence in child welfare practice*). Taylor jt (2019) analüüsisid samuti lastekaitsetöötajate juhtumitega seotud dokumente (sarnaselt ülal väljatoodule lastekaitsetöötajate juhtumikirjeldusi, juhtumiplaane, kohtudokumente ja uurimismenetlusega seotud andmeid), aga seda eesmärgiga välja selgitada, millist traumat on lastekaitsesüsteemis olevad lapsed kogenud ja kuidas lastekaitsetöötajad nende lastega töötavad. Larsson, Fredriksson ja Sjogren Fugl-Meyer (2019) seevastu analüüsisid haiglas töötavate sotsiaaltöötajate koostatud hinnanguid pikaajalise valu käes kannatavate patsientide sotsiaalsele ja psühho-sotsiaalsele toimetulekule. Kuna nimetatud tegurid mõjutavad olulisel määral patsientide heaolu, siis oli eesmärk vaadata, kas koostatud hinnangutes on vastavaid aspekte käsitletud, sest see omakorda mõjutab sekkumiste planeerimist patsientide rehabilitatsiooni plaani (Larsson, Fredriksson ja Sjogren Fugl-Meyer,

2019). Ühtlasi soovisid autorid teada, kas sotsiaaltöötajate hinnangud erinevad patsientide soo ja vanuse lõikes. Lisaks sellele kasutasid autorid oma uurimuses patsientide terviseandmeid.

Kõigis kolmes uurimuses lähtuti sotsiaaltöötajate hinnangute ja lastekaitsetöötajate juhtumite analüüsimisel vastavat valdkonda reguleerivatest või valdkonnas kasutusel olevatest juhendmaterjalidest, suunistest ja raamdokumentidest. See tähendab vaadati, kas spetsialistide töö (nii nagu nad seda juhtumitega seotud dokumentides kirjeldavad) on kooskõlas valdkondlikes dokumentides sätestatuga. Näiteks Carnochan jt (2018) ning Taylor jt (2019) lähtusid oma uurimustes California lastekaitsetöö raamistikust (ingl. k. *the Child Welfare Core Practice Model*), mis käsitleb lastekaitsetöö teoreetilisi lähtekohti, väärtusi, juhtumitöö komponente, lastekaitsetöötaja käitumise põhimõtteid jpm, mille eesmärk on toetada lastekaitsetöötajaid nende töös. Taylor jt (2019) lähtusid oma analüüsis täiendavalt veel trauma teadliku praktika n-ö juhistest (vastavalt eelnevalt välja toodud California lastekaitsetöö raamistiku soovitudele). Larssoni, Fredrikssoni ja Sjogren Fugl-Meyeri analüüs toetus rahvusvahelise funktsioneerimisvõime, vaeguste ja tervise klassifikatsioonile ehk RFK-le (täpsemalt vaadati, kas sotsiaaltöötajate hinnangud sisaldavad tegutsemise ja osaluse ning keskkonnateguritega seotud asjaolusid).

Nii Carnochan jt (2018) kui Taylor jt (2019) leidsid, et lastekaitsetöötajad lähtuvad oma töös ülal väljatoodud raamdokumentidest. Carnochan jt (2018) leidsid veel täiendavalt, et on tegevusi, mida juhendmaterjalid ei sisalda, kuid mida praktikas rakendatakse, nt konfliktivahendamine lapse ja hooldaja vahel. Seega võimaldab administratiivsete andmete analüüsimine ühest küljest saada teavet sotsiaaltöö, sh lastekaitsetöö praktika kohta ning teisest küljest saab praktika kohta saadud uut teadmist kasutada nt juhendmaterjalide, suuniste ja raamdokumentide täiendamiseks (Carnochan jt, 2018). Lisaks sellele näitlikustab Carnochan jt (2018) läbi viidud juhtumite analüüs seda, kuidas teooria ja praktika on omavahel seotud, seega leiavad autorid, et juhtumianalüüs peaks olema osa nii sotsiaaltöö koolitussüsteemist kui supervisioonist. Taylor jt (2019) leiavad omakorda, et lastekaitse administratiivsete andmete puhul on väärtuslik nende n-ö ajaline mõõde, st mitmed juhtumid, mida nad analüüsisid olid olnud aastaid aktiivsed.

Erinevalt eelmises lõigus käsitletud uurimustele leidsid Larsson, Fredriksson ja Sjogren Fugl-Meyer (2019) oma analüüsis, et sotsiaaltöötajad ei käsitle oma hinnangutes kõiki tegutsemise ja osaluse ning keskkonnateguritega seotud asjaolusid, nii nagu need RFK-s kajastuvad. Samas nad toovad välja, et kuna sotsiaaltöötajad töötavad multidistsiplinaarses meeskonnas, siis on

võimalik, et sotsiaaltöötajate poolt mitte hinnatud teemasid on hinnanud mõni teine meeskonnaliige. Siit võib järeldada, et alati ei piisa ühe spetsialistide grupi, aga ka laiemalt ühe valdkonna andmete analüüsimisest, sest tervikpildi saamiseks on oluline laiem n-ö valdkondade ülene vaade.

3.1.3 Sotsiaaltöö praktika edendamine

Gillingham ja Graham, 2016; Zetino ja Mendoza, 2019 ja Zhi-Ping, Zhi-Fei ja Yu-Zhou, 2020 räägivad suurandmete kasutamisest sotsiaaltöö praktika edendamise kontekstis. Kui Zetino ja Mendoza (2020) ning Gillingham ja Graham (2016) üksnes kirjeldavad võimalikke suurandmete kasutamise võimalusi ja sellega seotud väljakutseid, siis Zhi Ping, Zhi-Fei ja Yu-Zhou (2020) viisid läbi uurimuse, kus nad kasutasid suurandmeid, et välja selgitada, mis mõjutab eakate sotsiaalteenuste kasutamist ühes konkreetses Hiina linnas.

Suurandmete kasutamist sotsiaaltöö praktika edendamiseks on soodustanud erinevad arengud. Kõigepealt, mitmetes valdkondades on suurandmetel põhinevat andmeanalüüsi edukalt kasutatud, seda muuhulgas selleks, et inimeste vajaduste, soovide ja käitumise kohta mitmekülgsemat infot saada, et selle põhjal tõhusamalt ressursse planeerida, teenuseid ja tooteid arendada ning suunata (Gillingham, 2016; Zetino ja Mendoza, 2019). Siit edasi, kui suurandmeid analüüsides ning selle põhjal saadud teadmisi kasutades on olnud võimalik omavahel kokku viia näiteks inimesed ja tooted, siis sarnaselt võiks see omavahel kokku viia ka näiteks inimesed ja sotsiaalteenused, seda „kiiremini, nutikamalt ja efektiivsemalt“ (Zetino ja Mendoza, 2019: 410).

Elektrooniliste andmebaaside ja infosüsteemide kasutuselevõtt sotsiaaltöös on parandanud sotsiaaltöö agentuuride võimet teenuste ja teenuste saajatega seotud infot dokumenteerida, analüüsida ja raporteerida (Gillingham, 2016). Ühtlasi on tehnoloogia areng teinud võimalikuks erinevate andmebaaside koostoime (Gillingham, 2016; Zetini ja Mendoza, 2019), mis võimaldab saada inimesest ja tema olukorrast tervikliku ülevaate. Need arengud pakuvad täiendavaid võimalusi sotsiaaltöö tulemuslikkuse hindamiseks. Näiteks Zetino ja Mendoza (2019) kirjutavad, et teenuste ja sekkumiste mõju inimesele võib väljenduda erinevates

eluvaldkondades (sh hariduses, tervises, eluasemes jne) ning väga pika aja jooksul. Sealjuures mängivad olulist rolli inimese enda motivatsioon, tema minevik ja sotsiaalse tugivõrgustiku olemasolu. Kui aga asetada tulemused erinevates andmebaasides olemasoleva info konteksti, siis on võimalik analüüsida inimeste, teenuse pakkujate, sekkumiste/ravi ja tulemuste vahelisi seoseid ja nende alusel sekkumisi arendada (Zetino ja Mendoza, 2019). Lisaks sellele võimaldab selline tervikülevaade sarnaste tunnuste ja vajadustega inimeste profile luua (Zetino ja Mendoza, 2019) ning seeläbi sotsiaalteenuseid paremini „planeerida, pakkuja ja suunata“ (Gillingham, 2016:1).

Suurandmete potentsiaali ärakasutamiseks sotsiaaltöö praktika edendamiseks on muuhulgas oluline, et sotsiaaltöötajad oleksid teadlikud nii suurandmete kasutusvõimalustest kui suurandmete kasutamisega seotud väljakutsetest (Zetino ja Mendoza, 2019; Gillingham 2016). Zetino ja Mendoza (2019) ning Gillinghami (2016) kirjeldatud praktilised näited võiksid selles osas inspiratsiooni anda. Näiteks on andmeanalüütika pakutavad võimalused rakendust leidnud erinevates valdkondades (Zetino ja Mendoza, 2019). Kirjeldavat analüüsi saab kasutada näiteks n-ö kuumuskaartide loomiseks (ingl. k. *heatmap*), et visualiseerida huvi pakkuva nähtuse esinemist konkreetses piirkonnas. See on laialdast kasutust leidnud politseis, et kaardistada kuritegevuse esinemist erinevates piirkondades. Lisades sellele info erinevate piirkonda iseloomustavate näitajate kohta (nt teenuste kättesaadavus, infrastruktuur jms) on võimalik erinevate näitajate vahelisi võimalikke seoseid tuvastada (Zetino ja Mendoza, 2019). Ennustava analüüsi rakendusena toovad Zetino ja Mendoza (2019) välja perevägivalda esinemist ennustava mudeli. Täiendava näitena saab välja tuua Gillinghami (2016) kirjeldatud otsuste tegemist toetavad abivahendid, mida kirjeldasin põhjalikumalt ülal. Diagnostilist analüüsi saab Zetino ja Mendoza (2019) hinnangul kasutada näiteks vaimse tervise probleemidest parema ülevaate saamiseks ning ettekirjutavat analüüsi selleks, et sekkumisi koordineeritult planeerida.

Zhi-Ping, Zhi-Fei ja Yu-Zhou (2020) uurimus on läbi viidud Hiinas Guilini linnas, kus nad analüüsisid suurandmeid, et välja selgitada, millised tingimused mõjutavad eakate sotsiaalteenuste kasutamist. Autorid lähtusid oma uurimuses Anderseni käitumismudelilt, mis käsitleb tervishoiuteenuste kasutamist mõjutavaid tegureid. Nimetatud mudeli kohaselt on inimeste teenuste kasutamine seotud eelsoodumusega (demograafilised ja sotsiaalsed tegurid ning isiklikud tõekspidamised), võimaldavate ressurssidega (perekindlikud ja ühiskondlikud tegurid) ning vajadusega (subjektiivsed ja objektiivsed vajadused) (Zhi-Ping, Zhi-Fei ja Yu-Zhou, 2020).

Zhi-Ping, Zhi-Fei ja Yu-Zhou (2020) on oma uurimuses eelsoodumusega seotud teguriteks valinud inimese vanuse ja teenuste süsteemi, võimaldavateks ressurssideks igakuise sissetuleku ja perekonna suuruse (st perekonnaliikmete arvu, kes eakaga koos elavad) ning vajadustega seotud teguriteks meditsiiniliselt hinnatud sümptomid ja diagnoosi. Autorid uurisid, milline on väljatoodud tegurite ja teenuste kasutamise (sh teenuste kasutamise sageduse) vaheline seos. Selleks kasutasid nad sotsiaaltöoga seotud avalike platvormide ning valitsusasutuste andmebaaside andmeid ja erinevaid uuringuid ning analüüse. Täpsemaid detaile seoses kasutatud andmetega ei ole uurimuses välja toodud. Kokkuvõtvalt, Zhi-Ping, Zhi-Fei ja Yu-Zhou (2020) kirjutavad, et mida vanemad on inimesed, seda parem on nende teadlikkus sotsiaalteenuste olemasolust ja seda rohkem nad teenuseid kasutavad. Nad lisavad, et samuti kasutavad teenuseid rohkem need inimesed, kellel on raskusi igapäeva toimingute teostamisel. Lisaks sellele kasutavad teenuseid rohkem eakad, kes sõltuvad elamiskulude katmisel riiklikust pensionist ja teistest riiklikest toetustest ning vähem need eakad, kellel on muu individuaalne sissetulek (Zhi-Ping, Zhi-Fei ja Yu-Zhou, 2020). Autorid teevad uurimuse tulemustele tuginedes soovitusi sotsiaalteenuste edendamiseks.

3.1.4 Sotsiaalpoliitika kujundamine

Kolmes allikas (Kum, Stewart, Rose ja Duncan, 2015; Mader, 2015; Zhi-Ping, Zhi-Fei ja Yu-Zhou, 2020) kirjeldatakse suurandmete kasutamist, et toetada tõenditele tugineva sotsiaalpoliitika kujundamist ja elluviimist. Zhi-Ping, Zhi-Fei ja Yu-Zhou (2020) uuringut käsitlesin eelmises alapeatükis, aga kuna autorid tegid muuhulgas soovitusi poliitikakujundamiseks, siis on asjakohane ka käesolevas peatükis seda allikat kasutada. Teised autorid, Kum jt (2015) ning Mader (2015) kirjutavad administratiivsete andmete kasutamisest ja analüüsimisest eesmärgiga informeerida sotsiaalpoliitika. Autorite lähenemised on aga erinevad, st kui Kum jt (2015) kirjutavad teadmushõivest andmebaasidest (ingl *knowledge discovery in databases*) ja selle protsessist ning rakendamisest lastekaitse valdkonnas, siis Mader (2015) kirjutab n-ö traditsiooniliste uurimismeetodite kasutamisest ja nende kombineerimisest administratiivsete andmete analüüsimisel.

Kum jt (2015: 130) selgitavad, et teadmushõive andmebaasidest on „mittetriviaalne protsess, et leida andmetest paikapidavaid, uudseid, potentsiaalselt kasulikke ning kokkuvõtvalt

mõistetavaid mustreid“. See hõlmab endas andmete ettevalmistamist, puhastamist ja kombineerimist; muutujate ja tunnuste valikut; andmete analüüsimist, andmekaevandamist ja mudelite loomist; tulemuste valideerimist ja hindamist ning tulemustele reageerimist, tulemuste põhjal otsuste tegemist (Kum jt, 2015). Kum jt (2015) kirjeldavad, kuidas teadmushõive protsessi on rakendatud Ameerika Ühendriikides Põhja-Carolina osariigis, et toetada sotsiaalpoliitika kujundamist ja praktikat. Autorid selgitavad, et sellist andmepõhist lähenemist on Põhja-Carolinas kasutatud juba alates 1997. aastast, seda erinevate sotsiaalprogrammide (nt *Work First/Temporary Assistance for Needy Families* ehk TANF, *Food and Nutrition Services/Food Stamps*) elluviimisel ning lastekaitse valdkonnas. Autorid keskenduvad viimasele.

Kui lühidalt kokku võtta, siis Kum jt (2015) kirjeldavad Põhja-Carolina ülikooli ja Põhja-Carolina sotsiaalteenuste osakonna (ingl *Division of Social Services*) koostööd teadmushõive protsessi rakendamisel lastekaitse süsteemi administratiivsetele andmetele, mille eesmärk on anda põhjalik ülevaade osariigi lastekaitse süsteemist. Autorid selgitavad, et selle koostöö tulemusel on loodud dünaamiline veebileht, mis võimaldab saada nii osariigi ülest kui kohalike omavalitsuste põhist ülevaadet/infot erinevate lastekaitsetöö kontekstis oluliste näitajate kohta. Seega saab seda kasutada erinevate ülevaadete koostamiseks, muutuste jälgimiseks, otsuste tegemiseks jne (Kum, 2019). Lisaks sellele kasutatakse neid andmeid erinevate analüüside, uuringute ja raportite koostamisel, sh saab neid värske andmetega pidevalt uuendada. Autorid lisavad, et loodud lahendus on samuti väga kasulik erinevate teadustööde tegemisel ja toovad sellekohaseid näiteid.

Kum jt (2015) lisavad, et teadmushõive töötab kõige paremini siis, kui andmeid loovad ja nende andmete analüüsimise põhjal saadud teadmisi kasutavad n-ö ühed ja samad inimesed. Autorite hinnangul aitab see tagada andmete kvaliteeti ja seda, et andmeid tõlgendatakse õigesti. Ühtlasi peavad loodud rakendused (antud juhul veebileht) erinevate sihtrühmade jaoks kasulikud olema, autorite uurimuses sotsiaalprogrammide haldajatele ja lastekaitse valdkonnas töötavatele spetsialistidele.

Valitud lähenemise, st riigiasutuste ja avalike ülikoolide koostöös teadmushõive süsteemi väljatöötamise osas toovad Kum jt (2015) välja, et see on ennast mitmel põhjusel õigustanud. Näiteks nad selgitavad, et see on hea võimalus praktika ja uurimuse seostamiseks, kus praktikud saavad süsteemi loojatele anda infot oluliste lastekaitse valdkonda puudutavate poliitikate, praktikate ja konteksti puudutavate tegurite kohta ning ülikooli meeskond saab

omakorda selle põhjal uusi lahendusi välja töötada, rääkimata erinevate uuringute ja poliitikanalüüside koostamisest ning uute lähenemiste väljatöötamisest. Autorid lisavad, et sellise koostöö lisaväärtus seisneb veel selles, et ülikoolidel on võimalik interdistsiplinaarseid meeskondi moodustada, kuhu kuuluvad nii sisuekspertid (nt sotsiaaltöö, avaliku halduse, politoloogia ekspertid) kui arvutiteaduse valdkonna ekspertid. Kuna lastekaitse valdkonna asutustel selliseid võimalusi ei ole, siis see tähendab, et valitud lähenemisega on võimalik andmesüsteemi potentsiaali täiel määral ära kasutada (Kum jt, 2015).

Sarnaselt Kum jt (2015) rõhutab Mader (2015) administratiivsete andmete olulisust poliitika kujundamisel. Ta selgitab, et administratiivsed andmed, mis sisaldavad inimeste kohta infot n-ö sünnist-surmani on sotsiaalvaldkonnas väärtuslik ressurss. Sealjuures loob andmete paranev kvaliteet, ulatus ja andmebaaside koostoime uusi võimalusi neid andmeid erinevate uurimismeetodite kombineerimisega analüüsida (Mader, 2015). Lisaks sellele kirjutab Mader (2015), et administratiivsete andmete näol on uurijate käsutuses samasugune info (inimeste, programmide jms kohta), mis poliitikakujundajatel erinevate poliitikate väljatöötamisel, mistõttu on nende analüüsimine poliitikate informeerimiseks eriti asjakohane.

Konkreetselt, Mader (2015) näeb, et omavahel ühendatud administratiivsed andmesüsteemid on uurimuslikus mõttes väärtuslikud mitmel põhjusel. Lisaks sellele, et valimid on suured ja andmeid luuakse pidevalt juurde ning erinevate asutuste poolt pakutud teenused on mitmekesised (sh on nende loomisel lähtunud sotsiaalpoliitika prioriteetidest), siis nagu eelnevalt mainitud, “jälgivad“ need inimesi pika aja jooksul, mistõttu saab analüüsida erinevate tegevuste võimalikku mõju inimese elukäigule (Mader, 2015). Lisaks sellele annavad need infot teenuseid pakkuvate asutuste ja neis töötavate inimeste kohta, nt on võimalik lähemalt uurida neid organisatsioone, mis on oma tegevustes efektiivsed (Mader, 2015). Veel leiab Mader (2015), et omavahel ühendatud administratiivsed andmesüsteemid võimaldavad uurida, millised on ühesuguste tegevuste tulemused erinevate inimeste jaoks (sh võttes arvesse teisi tegureid) ning samuti on võimalik uuringuid korrata (nt võrrelda erinevaid sarnaste tunnustega grupe erinevatel ajahetkedel, et analüüsida erinevate poliitikates tehtud muudatuste tagajärgi inimeste jaoks). Täiendavalt kirjeldab Mader (2015) administratiivsete andmete kasutamisega seotud väljakutseid, kuid neid käsitlen põhjalikumalt järgmises peatükis.

3.2 Suurandmete kasutamisega seotud väljakutsed kirjanduse põhjal

Analüüsitud kirjanduse põhjal on suurandmete kasutamisega seotud väljakutsed sotsiaaltöös seotud valdkonnas kasutatavate andmebaaside (Gillingham ja Graham, 2017; 2019a; 2019b, 2020; Keddell, 2019; Lery, Haight ja Alpert, 2016) ja infosüsteemidega (Gillingham ja Graham, 2017; 2019a, 2019b; Kum, Stewart, Rose ja Duncan, 2015; Schneider ja Seelmeyer, 2019) ning sealt edasi andmetega, mida dokumenteeritakse, kogutakse ja analüüsitakse (Carnochan, Weissinger, Henry, Liner-Jigamian ja Austin, 2019; Gillingham ja Graham, 2017, 2019a; 2019b; 2020; Mader, 2015; Schneider ja Seelmeyer, 2019; Taylor, Battis, Carnochan, Henry, Balk ja Austin, 2019). Lisaks sellele saab väljakutsena välja tuua erinevaid eetilisi küsimusi, mis seoses suurandmete kasutamisega tõusetuvad (Keddell, 2015; Gillingham ja Graham, 2017; 2019a; 2019b; Mader, 2015; Schneider ja Seelmeyer, 2019; Zetino ja Mendoza, 2019). Kuivõrd nimetatud teemad on omavahel seotud, siis käsitlen neid käesolevas peatükis koos.

Elektroonilised infosüsteemid ja andmebaasid on teinud võimalikuks erinevate andmete kogumise sotsiaaltöös, sh tunduvalt suuremas mahu, kui see varasemalt võimalik oli (Gillingham ja Graham, 2016). Samas, nagu Lery jt (2016) kirjutavad, siis andmebaasid ei ole väärtus omaette, vaid väärtust saab luua, kui neis sisalduvaid andmeid oskuslikult analüüsida. Seda toetavad tehnoloogia areng, sh üha võimsamad arvutid ning andmete kogumine nii teenuste kasutajate kui teenuste pakkumise kohta koos võimalusega erinevate andmebaaside infot kasutada (Gillingham, 2019a). Viimase osas on oluline välja tuua, et andmebaaside koostoime võimalused on riigiti erinevad (Keddell, 2019), ning väljakutseks võib olla nii asutuste vaheline andmebaaside koostoime kui ühe asutuse sisene infovahetus asutuse erinevate osade vahel (Penuel ja Means, 2011, viidatud Kum jt, 2015). Seega saab kirjanduse põhjal öelda, et suurandmete kasutamisel sotsiaaltöös peab arvestama mitme asjaoludega.

Gillingham (2020) kirjutab, et elektrooniliste infosüsteemide kasutuselevõtt on muutunud, mida sotsiaaltöötajad dokumenteerivad ning kuidas nad seda teevad või teha saavad. Infosüsteemide ülesehitus mõjutab „kuidas praktikud mõtlevad, millele tähelepanu pööravad ja mida kirja panevad ning mida mitte“ (Gillingham 2016, viidatud Gillingham 2020). Seega hästi disainitud lahendused võimaldavad andmeid parimal võimalikul moel kasutada (Lery jt, 2016). See, milliseid võimalusi infosüsteem dokumenteerimiseks pakub sõltub paljuski infosüsteemi loojatest (Gillingham, 2020). Tuginedes Collingridgele (1980) kirjutavad Schneider ja

Seelmeyer (2019), et juba tehnoloogia planeerimise, arendamise ja disainimise etapis on oluline tehnoloogiat hinnata, et vältida selle hilisemat negatiivset mõju.

Lisaks sellele on oluline, et ollakse teadlik erinevatest andmetega seotud asjaoludest, sest sotsiaaltöös loodavate andmete kvaliteet on oluline väljakutse suurandmete kasutamisele sotsiaaltöös (Mader, 2015). Andmeid loovad inimesed, kes nagu eelnevalt öeldud otsustavad, mida dokumenteerida ja mida mitte (Gillingham ja Graham, 2016). Taylor jt (2019) ning Carnochan jt (2019), kes uurisid lastekaitsetöötajate märkmeid kirjutavad, et spetsialistide vahel on selles osas olulisi erinevusi (Taylor jt, 2019), sealjuures on võimalik, et oma tegevusega seoses kajastatakse rohkem positiivseid kui negatiivseid aspekte (Carnochan jt, 2019). Täiendavalt, Gillingham ja Graham (2016), tuginedes Huuskonen ja Vakkari (2013) kirjutavad, et sotsiaaltöötajad lähtuvad dokumenteerimisel oma pragmaatilistest vajadustest, mis ei pruugi olla kooskõlas infosüsteemi loojate kavatsustega. Andmebaasides kajastub n-ö ametlik versioon praktikast ja see võib olulisel määral erineda n-ö mitteametlikust versioonist (Gillingham, 2019a). Gillingham (2020), viidates lastekaitse süsteemi andmetele kirjutab, et selle põhjuseks on juhtumite dokumenteerimise eesmärk. Ta lisab, et lastekaitse süsteemis on dokumenteerimise eesmärk tõendada, et lastekaitsetöötaja reageeris olukorrale, tegi sellega seoses otsuseid ning tegi neid vastavalt seadustele. Seega ei anna andmed täit ülevaadet sotsiaalsetest probleemidest ega sellest, kuidas teenuste kasutajad ise oma elu kogevad (Gillingham, 2020).

Suurandmete analüüsimiseks kasutatakse erinevaid tehnikaid ning analüüsitud kirjanduses on palju käsitlust leidnud masinõppe võimalused (Gillingham ja Graham, 2016, 2019, a, 2019b, 2020; Carnochan jt, 2019; Keddell, 2019; Kum jt, 2017; Larsson, Fredricksson ja Sjogren Fugl-Meyer, 2019; Schneider ja Seelmeyer, 2019; Shwartz jt, 2017; Taylor jt, 2019; Zetino ja Mendoza, 2019). Tõsi, masinõppe meetoditega on võimalik analüüsida rohkem andmeid kui inimesed seda suudavad (Keddell, 2019), kuid andmete suur maht ei tähenda ilmingimata, et valim, mille põhjal järeldusi ja soovitusi tehakse on esinduslik (Gillingham ja Graham, 2016). Gillingham ja Graham (2016) kirjutavad, et sotsiaaltöös kogutud andmed annavad infot üksnes nende inimeste kohta, kellel on süsteemiga kokkupuude olnud. Kusjuures erateenuseid kasutavate inimeste andmed administratiivsete andmete hulgas ei kajastu (Mader, 2015). Mõlemad, Gillingham (2020) ja Keddell (2019) lisavad veel, et näiteks lastekaitse süsteemi andmed ei kajasta laste väärkohtlemist kogu selle ulatuses, vaid üksnes seda osa, mis jõuab lastekaitse süsteemi tähelepanu alla. Täiendavaks probleemiks on, et lastekaitse süsteemis on teatud grupid n-ö üle esindatud (nt USA kontekstis rassilised ja etnilised vähemused),

alaesindatud (nt hea sotsiaalmajandusliku taustaga perekonnad) ja lisaks on veel selliseid gruppe, kelle info ei pruugi andmetes üldse kajastuda (Keddell, 2019). Kui neid andmeid kasutada otsustustoe süsteemide loomisel, siis võib see viia kindlate tunnustega inimeste juhtumitele kas üle- või alareageerimisele (Keddell, 2019; Gillingham 2019b). Tekib oht, et teatud ühiskonnagrupid langevad diskrimineerimise ja ebavõrdse kohtlemise ohvriks (Keddell, 2019). See, et kindla grupi sarnasuste põhjal tehakse otsuseid individuaalsete juhtumite kohta tähendab omakorda, et inimeste olukordadele ei läheneta individuaalselt (Keddell, 2019).

Gillingham ja Graham (2017) kirjutavad, et suurandmete kasutamisel räägitakse andmete objektiivsusest. Näiteks Zetino ja Mendoza (2019) kirjutavad, et kui inimese elu n-ö vaadata erinevate andmebaaside kontekstis, siis see annab inimesest täielikuma ja objektiivsema ülevaate. Samas Keddell (2019) leiab, et selline arusaam viib olukorrani, kus inimeste enda arvamuste ja arusaamadega oma elust ei arvestata, sh jäävad tähelepanuta teised kaitsvad ja riskitegurid, mis andmetes ei kajastu. Mader (2015) lisab, et andmetes mittekajastuva info teadvustamine on muuhulgas oluline olukorras, kus andmete põhjal tehakse järeldusi erinevate sekkumiste efektiivsuse kohta, sest see sõltub inimeste individuaalsetest omadustest, spetsialistidest jms. Seega, sotsiaaltöös kogutakse küll väga palju erinevaid andmeid (kvantitatiivseid, kvalitatiivseid, struktureeritud, struktureerimata), aga üksnes andmete maht ei taga andmete objektiivsust ja neutraalsust (Schneider ja Seelmeyer, 2019). Gillingham ja Graham (2016), tuginedes Crawfordile (2013) selgitavad, et eelnevast tulenevalt ei asenda suurandmed täielikult kvalitatiivseid meetodeid, sest isegi kui suurandmete analüüsimisega tuvastatakse uusi mustreid, siis on sotsiaaltöös endiselt oluline sisuliselt aru saada, mis nende mustrite taga peitub.

Keddell (2019) rõhutab, et sotsiaaltöötajatel on tulenevalt sotsiaaltöö väärtustest ja eetilistest tõekspidamistest oluline roll suurandmete rakendamisel valdkonnas kaasa rääkida. Nagu eelnevast väljatoodust selgus, siis on mitmeid andmebaaside ja infosüsteemide ning andmete olemuse ja kvaliteediga seotud asjaolusid, mis võivad viia andmepõhise lähenemise ja sotsiaaltöös oluliste põhimõtete vastuoluni. Lisaks eelnevast väljatoodust on üheks peamiseks eetikaga seotud aruteluteemaks inimeste privaatsus (Gillingham ja Graham, 2016; Keddell, 2019; Schneider ja Seelmeyer, 2019; Zetino ja Mendoza, 2019). Näiteks Gillingham ja Graham (2017) kirjutavad teadliku nõusolekuga seoses, et inimestel on seda keeruline anda, sest sellel hetkel, kui andmeid kogutakse ei teata, mil moel neid hiljem kasutatakse.

4 ARUTELU

Selles peatükis arutlen käesoleva magistritöö analüüsitulemuste üle. Minu analüüsitud teadusartiklites käsitletakse sotsiaaltöö valdkonna suurandmetena infosüsteemides ja andmebaasides talletatud administratiivseid andmeid. Uurimustöö valimi moodustanud allikates kirjeldatakse administratiivsete andmete, sh andmete loomise, kogumise, analüüsimise ja kasutamisega seotud võimalusi ja väljakutseid. Teiste andmete, st interneti rakenduste ja nutistu seadmete kasutamise tulemusel tekkivate andmete kasutamisest sotsiaaltöö valdkonnas analüüsitud teadusartiklites ei kirjutata.

Analüüsitulemustest ilmneb, et sotsiaaltöö valdkonnas on suurandmeid kasutatud sotsiaalpoliitika kujundamiseks, sotsiaaltöö praktika uurimiseks ja edendamiseks ning otsustustoe süsteemide loomiseks. Kuigi selguse mõttes käsitlesin neid teemasid analüüsi peatükis eraldiseisvate kategooriatena, siis on need omavahel selgelt seotud.

Olenemata suurandmete kasutamise võimalustest sotsiaaltöös on oluline mõista sotsiaaltöö valdkonna suurandmete olemust, sest suurandmete rakendusvõimalused ja nende oodatav lisaväärtus sõltuvad paljuski just andmetest, millele need tuginevad. Sellest tulenevalt keskendun käesolevas peatükis sotsiaaltöö valdkonna suurandmetele (andmete loomise, kogumise, analüüsimise ja kasutamisega seotud võimalustele ja väljakutsetele), andmeteadlikkuse olulisusele ja andmete loomise oskusele sotsiaaltöös ning suurandmete kasutamisega seotud eetilistele küsimustele mitte niivõrd konkreetsetele suurandmete rakendusvõimalustele, mida kirjeldasin käesoleva magistritöö analüüsi peatükis.

4.1 Suurandmed sotsiaaltöös

Nagu öeldud, siis käesolevas magistritöös analüüsitud artiklites käsitletakse sotsiaaltöö valdkonna suurandmetena infosüsteemides ja andmebaasides talletatud andmeid. Ühtlasi selgus minu uurimustööst, et kogutud andmeid on võimalik mitmel erineval moel kasutada, kuid seda tehes on oluline teadvustada erinevaid andmete olemusega seotud aspekte.

Analüüsitud teadusartiklite põhjal saa öelda, et eeldus suurandmete kasutamiseks sotsiaaltöö valdkonnas on olemas, st infosüsteemid ja andmebaasid on sotsiaaltöös laialdaselt kasutusel ning sellest tulenevalt loovad sotsiaaltöötajad oma igapäevatöö raames palju digitaalseid andmeid, mida on võimalik analüüsida. Lisaks sellele selgus magistritöös analüüsitud artiklitest, et andmed, mida sotsiaaltöötajad loovad ja info, mida nad koguvad ning dokumenteerivad on väga mitmekülgne.

Näiteks Schneider ja Seelmeyer (2019) kirjutavad, et sotsiaaltöös on nii kvantitatiivseid kui kvalitatiivseid andmeid, struktureeritud ja struktureerimata andmeid, milles kajastub nii teenustega kui teenuste saajatega seotud info. Taylor jt (2019) ning Carnochan jt (2019) uurimused ilmestavad lastekaitsetöös loodavate ja kogutavate andmete mitmekesisust, mille hulgas on laste ja perekondade olukorda kirjeldavad andmed, info sekkumiste planeerimise, elluviimise ja tulemuste kohta, kohtu dokumendid jpm. Larsson jt (2019) uurimus omakorda näitab, kui mitmekesised võivad olla haiglas töötavate sotsiaaltöötajate poolt loodud ja kogutud andmed. Ka laiemalt teaduskirjanduse põhjal saab öelda, et sotsiaaltöös luuakse ja kogutakse väga palju erinevaid andmeid. Näiteks Reamer (2005) kirjutab, kuidas sotsiaaltöö valdkonna dokumentatsioonis kajastub info sotsiaaltöötajate arutluskäigu, hinnangute ja tegevuse kohta seoses klienditööga, sekkumiste planeerimise ja elluviimisega, koostööga jpm.

Seega kui vaadata sotsiaaltöös loodavate andmete mitmekesisust, siis tundub, et need justkui annavad erilise sissevaate nende inimeste maailma, kellega sotsiaaltöötajad töötavad. Sealhulgas sotsiaalsestesse probleemidesse, inimeste nõrkustesse ja tugevustesse. Lisaks sellele ka sotsiaaltöötajate praktikasse, teadmistesse ja oskustesse. Sellega seoses tõusetus käesolevas magistritöös analüüsitud teadusartiklitest aruteluteema seoses sotsiaaltöös loodud ja kogutud andmete objektiivsuse ja erapooletusega. Schneider ja Seelmeyer (2019) kirjutavad, kuidas suurt hulka mitmekesiseid andmeid võidakse seostada sellega, et need kirjeldavad maailma objektiivselt ja erapooletult. Teised autorid (Gillingham ja Graham, 2017; Keddell, 2019; Gillingham 2019a, 2019b, 2020, Lery jt, 2016) seevastu selgitavad, millisel põhjusel see sotsiaaltöö valdkonnas nii ei ole. Näiteks Gillingham ja Graham (2017) ning Gillingham (2020) kirjeldavad, kuidas sotsiaalteenuste saajate kohta talletatud andmed võivad olla ebatäpsed ja ebatäielikud tulenevalt sellest, et spetsialistid sisestavad andmeid andmebaasi kas alles tööpäeva lõpus või isegi hiljem. Lisaks sellele võivad sotsiaaltöö valdkonna andmed olla kallutatud (Gillingham ja Graham, 2017; Keddell; 2019). Seda nii seetõttu, et sotsiaaltöötajad tunnevad seoses dokumenteerimisega survet asju kindlat moodi kirja panna (nt kui tajutakse,

et kirja pandut kasutatakse spetsialisti töö kontrollimiseks) (Gillingham ja Graham, 2017) või seetõttu, et teatud ühiskonnagrupid on sotsiaaltöös n-ö ülesindatud (Keddell, 2019).

Sellised andmete kvaliteedi ja esinduslikkusega seotud probleemid ei ole küll sotsiaaltöö valdkonna spetsiifilised, kuid neil on sotsiaaltöö kontekstis eriline tähendus. Nagu minu uurimustööst selgus, siis sotsiaaltöö valdkonna suurandmeid saab kasutada erinevatel eesmärkidel. Andmete põhjal tehakse olulisi inimeste elu mõjutavaid otsuseid ning valede otsustel võivad olla väga traagilised tagajärjed. Samas kui andmed justkui räägivad iseenda eest ja andmete tõlgendamisel ei arvestata laiemat kontekstiga, siis on keeruline andmete põhjal tehtud otsuseid vaidlustada. Nii minu magistrisöö analüüsitud artiklites (Keddell, 2019) kui teaduskirjanduses laiemalt (Gillingham, 2011; Devlieghere, Roose ja Evans, 2020) räägitakse seega sotsiaaltöötaja valdkondliku ekspertiisi ja kaalutusõiguse olulisusest, et lisaks andmetele oleks sotsiaaltöötajatel otsuste tegemisel võimalik arvestada ka teiste andmetes mittekajastuvate asjaoludega. Gillingham (2011), tuginedes Munrole (2005) kirjutab, et teoreetiliste teadmiste kõrval kasutavad sotsiaaltöötajad otsuste tegemisel ka n-ö elutarkust, neid kogemusi ja teadmisi, mida nad on elu jooksul omandanud ja mida andmetes ei ole.

Lisaks eelnevalt väljatoodule on suurandmete kasutamisel sotsiaaltöös oluline teadvustada ka valdkonnas kasutusel olevate infosüsteemide ja andmebaasidega seonduvat. Teaduskirjanduses laiemalt kirjutatakse infosüsteemide ja andmebaaside eelistest sotsiaaltöös. Infosüsteemide ja andmebaaside kasutamine sotsiaaltöös on üks võimalus sotsiaaltöö ja sotsiaaltöötajate tegevuse läbipaistvust ja vastutust suurendada, seda erinevate sihtrühmade ees alates teenuse saajatest kuni laiemale avalikkuseni (Devlieghere jt, 2020; Parton, 2008). Andmebaaside kindel struktuur toetab järjepidevust dokumenteerimisel ning aitab ühtlustada dokumenteerimise praktikaid (Devlieghere, jt 2020). Juhtumikirjeldused, hinnangud ja muu vabatekstiväljadel kajastatud info annab väärtuslikku teadmist nii inimeste (Coulton jt, 2015) kui sotsiaaltöö praktika kohta (Taylor, jt, 2019; Larsson jt, 2019; Carnochan jt, 2019). Samuti võimaldavad infosüsteemid ja andmebaasid teiste asutustega infot vahetada, erinevaid ühiskonnas toimuvaid protsesse paremini mõista ja probleemkohti tuvastada ning sotsiaaltöötajal oma tavapärasele tegutsemisele alternatiivseid lähenemisi kaaluda (Devlieghere, 2017; Scannapieco ja Iannone, 2014; viidatud Devlieghere, Roose ja Evans, 2020). Samas, väljatoodud eeliste taustal on probleemid ja väljakutsed, mida on oluline teadvustada.

Infosüsteemid ja andmebaasid ei ole n-ö erapooletud tööriistad, mille abil on võimalik andmeid koguda ja jagada, vaid neid peab vaatama konkreetse asutuse kontekstis, kus nad panustavad

teadmiste loomisse, juhtimisse jne (Kitchin, 2014). Infosüsteemide ja andmebaaside loomisega on seotud erinevad sihtrühmad, kes kõik toovad sellesse protsessi kaasa oma maailmanägemuse ja arusaama sotsiaalsetest ja individuaalsetest probleemidest (Introna, 2007 viidatud Gillingham, 2020 järgi). See mõjutab infosüsteemide hilisemat ülesehitust ja sotsiaaltöötajate võimalusi oma tööd kirjeldada ning inimeste, sekkumiste ja sekkumiste tulemusel saavutatud muutustega seonduvat dokumenteerida. Nagu Gillingham (2020), viidates Gillinghamile (2016) kirjutab, siis infosüsteemide ülesehitus mõjutab „kuidas praktikud mõtlevad, millele tähelepanu pööravad ja mida kirja panevad ning mida mitte“. Seega infosüsteemid loovad n-ö raamistiku, mis rohkem või vähem mõjutab sotsiaaltöötaja tegevust, võimalik et isegi prioriteetide ja eesmärkide seadmist ning seda teadmist, mida sotsiaaltöötaja oma töös kasutab (Ley ja Seelmeyer, 2008; viidatud Gillingham ja Graham, 2016). Sellest tulenevalt ei ole ka sotsiaaltöös loodud andmed n-ö erapooletud, vaid andmeid peab tõlgendama selles kontekstis, kus nad on loodud.

Lisaks sellele leiavad Devlieghere, Roose ja Evans (2020), et sotsiaaltöö valdkonnas kasutusel olevate infosüsteemide oluliseks puuduseks on, et need liigselt piiravad sotsiaaltöötajate võimalusi klienditöö erinevate etappidega seonduvat kirjeldada ning inimeste vajadustele paindlikult reageerida. Nende hinnangul on suhetele tuginev praktika asendunud „linnukeste“ tegemisega „õigetesse“ kastikesse ning kõik, mis etteantud raamidesse ei mahu läheb kaduma. Andmebaasis kajastub seega n-ö ametlik versioon praktikast, mis võib olulisel määral erineda n-ö mitteametlikust ja/või tegelikust versioonist (Buckley 2003; viidatud Gillingham 2019a ja Gillingham, 2011). Samas, kasutades andmebaaside infot inimeste vajaduste hindamisel ja teenuste planeerimisel, kasutatakse just neid andmeid ja seda infot, mis on andmebaasi sisestatud ning mis võib olla kallutatud või erinev.

Kokkuvõtvalt, suurandmed ja üha paremad võimalused andmeid analüüsida ning analüüsi tulemusi läbi erinevate innovaatiliste lahenduste valdkonda n-ö tagasi suunata ei asenda sotsiaaltöötajat ja tema valdkondlikku ekspertiisi. Nagu eelnevalt öeldud, siis andmed ei pruugi anda täit ja õiget ülevaadet inimestest, nende probleemidest ja vajadustest. Andmed, mida sotsiaaltöös luuakse, kogutakse ja hiljem analüüsitakse ei ole erapooletud ja n-ö ei räägi iseenda eest, vaid vajavad lahti mõtestamist ning konkreetsesse konteksti paigutamist. Just siin on sotsiaaltöötajatel oluline roll. Seda näiteks siis, kui andmete põhjal luuakse otsustustoe süsteeme. Sotsiaaltöötajad saavad aidata lahti mõtestada, milliste järelduste tegemine andmete põhjal on asjakohane ja milliste tegemine mitte. Samas, selleks et sotsiaaltöötajad saaksid nendes aruteludes osaleda ja arengutes kaasa rääkida peavad neil olema vastavad pädevused.

4.2 Andmeteadlikkus sotsiaaltöös

Coulton jt (2015) kirjutavad, et suurandmete potentsiaali kasutamiseks sotsiaaltöö valdkonnas on vaja sotsiaaltöötajaid, kellel on hea andmekirjaoskus ja kes teadvustavad, et neil on andmete loomisel oluline roll. Samas, nagu Santiago ja Smith (2019) selgitavad, siis vaatamata sellele, et sotsiaaltöö valdkonnas töötavad inimesed (nii sotsiaaltöö valdkonna juhid kui sotsiaaltöötajad) peavad koguma, talletama ja analüüsima suurt mahtu klientidega seotud administratiivseid andmeid (sh tegema nende andmete põhjal teenuste elluviimise ning teenuste rahastamisega seotud otsuseid), ei ole neid selleks eraldi ette valmistatud.

Digitaalset kirjaoskust (The Digital..., i.a) ja andmekirjaoskust (Kukke, 2020) peetakse hädavajalikeks 21. sajandi oskuseks. Digipädevuse enesehindamiseraamistik 2.1 (i.a) annab ülevaate digipädevustest ja digipädevustega seotud teadmistest, oskustest ning hoiakutest. Raamistikus kirjeldatakse digipädevusi viie pädevusvaldkonna (info, kommunikatsiooni, sisuloome, ohutus ja probleemilahendus) ja 21 pädevuse lõikes (The Digital..., i.a). Raamistikku saab näiteks kasutada selleks, et iseenda digipädevust hinnata. Andmekirjaoskus seevastu on seotud andmete mõtestamise ja kasutamise oskusega (Kukke, 2020). See hõlmab endas selliseid oskusi nagu leidmisoskus, lugemisoskus, töötlemisoskus, analüüsioskus ning argumenteerimisoskus (Kukke, 2020). Lisaks sellele on teaduskirjanduses hakatud digitaalse- ja andmekirjaoskuse kõrval rääkima ka kriitilisest suurandmete kirjaoskusest (Sander, 2020). Seda just seetõttu, et, inimeste teadlikkust suurandmetest ja suurandmete kasutamise võimalustest ning sellega seotud ohtudest ja riskidest tõsta (Sander, 2020).

Nimetatud oskused on kahtlemata olulised sotsiaaltöö valdkonnas, kus nagu analüüsist selgus, on mitmeid võimalusi andmetest informeeritud sotsiaaltöö tegemiseks. Nagu magistritöö sissejuhatuses kirjutasin, siis tööjõuvajaduse seire- ja prognoosisüsteemi OSKA ülevaates valdkonnaspetsiifiliste IKT-oskuste vajaduse kohta selgitatakse, et olenemata majandus- ja eluvaldkonnast on järgmise viie aasta kontekstis olulised „teadmised andmeanalüüsi kontseptsioonidest ja põhimõtetest“ ning „suurandmete analüüsioskus“ (OSKA ülevaade..., 2020: 6). Sealjuures on sotsiaaltöö valdkonna juhtide, tippspetsialistide ning sotsiaaltöötajate puhul leitud, et neil on vaja oskust suurandmeid analüüsida, analüüsi tulemusi tõlgendada ja suurandmete kasutusvõimalusi tunda (OSKA ülevaade..., 2020). Lisaks suurandmete kasutusvõimaluste tundmisele ning analüüsimise oskusele on sotsiaaltöö valdkonna juhtidel, tippspetsialistidel ja sotsiaaltöötajatel vaja veel teisi (sh suurandmete kontekstis) olulisi oskusi nagu digiturvalisuse tagamine (andmete turvaline jagamine ja isikuandmete kaitsest

kinnipidamine), IKT võimalusi kasutades teenuste disainimine (sh e-teenuste loomine), oskus tellida vajaduspõhiseid tehnoloogilisi lahendusi (sh vastavalt eesmärgile sobiva lahenduse valimine) ning valdkonnaga seotud mobiilirakenduste programmeerimise ja disainimise oskus. Saab öelda, et ootused sotsiaaltöötajatele on kõrged.

Erinevad uurijad (Parker-Oliver ja Demiris, 2006; Nguyen, 2007; Naccharati, 2010) on leidnud, et sarnaselt meditsiinivaldkonnale, kus on võimalik õppida näiteks tervishoiuinformaatikat ja õenduse informaatikat võiks sotsiaaltöös olla võimalik spetsialiseeruda näiteks sotsiaaltöö informaatikale (ingl *social work informatics*) (Parker-Oliver ja Demiris, 2006) ja/või lastekaitse informaatikale (ingl *child welfare informatics*) (Nguyen, 2007; Naccharati, 2010). Parker-Oliver ja Demiris (2006) kirjeldavad sotsiaaltöö informaatiku rolli järgmiselt: 1) sotsiaaltöö informaatik on kui uurija, kes aitab hinnata sotsiaaltöö valdkonnas tegutsevate asutuste tehnoloogilist valmisolekut ja konkreetsete tehnoloogiliste lahenduste kasutamist, sh praktikute ja klientide rahulolu selliste vahenditega, 2) sotsiaaltöö informaatik on kui uuendaja, kes toetab innovatsiooni sotsiaaltöös ning teeb seda vastavalt sotsiaaltöö eetikale ja 3) sotsiaaltöö informaatik on kui õpetaja/koolitaja, kes toetab sotsiaaltöötajaid uute teadmiste omandamisel, et spetsialistid oskaksid tehnoloogilisi lahendusi parimal võimalikul moel kasutada.

Naccharati (2010) kirjelduse kohaselt täidaks lastekaitse informaatik sarnaseid rolle, aga ta on eraldiseisvalt tähelepanu juhtinud veel sellele, et lastekaitse informaatikul peaks olema pädevus toetada otsustustoe süsteemide rakendamist lastekaitsetöös. Käesolevas magistritöös analüüsitud teadusartiklitest selgus, et üks väga konkreetne suurandmete kasutamise võimalus sotsiaaltöös on luua suurandmetele tuginevaid otsustustoe süsteeme. Ühtlasi selgus, et otsustoe süsteemid on laialdaselt kasutusele võetud just lastekaitse süsteemis ning see võib olla ka põhjus, miks Naccharati (2010) rõhutab just otsustustoe süsteemide kasutamisega seotud pädevusi.

Seega on arutelukoht, kas sotsiaaltöö valdkonnas oleks vajalik ja eesmärgipärane uusi spetsialiseerumise võimalusi luua või mitte. Täna Eesti kontekstis ilmselt räägime pigem interdistsiplinaarsest koostööst sotsiaaltöötajate, arvuteadlaste, statistikute jt spetsialistide vahel, et suurandmete potentsiaali sotsiaaltöö valdkonna jaoks ära kasutada.

4.3 Sotsiaaltöö eetika ja suurandmed

Sotsiaaltöö professiooni üheks tunnuseks on erialased eetikakoodeksid, mis nagu Selg (2010:22) kirjutab „põhinevad humanistlikel ja professionaalsetel väärtustel ning nende mõte on anda pidepunkte mõtisklusteks, kahtlustega toimetulekuks ja mitmekülselt kaalutletud otsusteks, kuidas on õigem ühes või teises olukorras toimida“. Sotsiaaltöö valdkonna eetikakoodeksid ja -juhised on ajas muutunud ning Reameri (2017) hinnangul on sotsiaaltöö eetika jõudnud n-ö digitaalsesse perioodi, kus tähelepanu keskmes on sotsiaaltöötajate info- ja kommunikatsioonitehnoloogia kasutamisega seotud eetilised küsimused sotsiaaltöös.

Kuigi Eesti sotsiaalala töötaja eetikakoodeksis (Sotsiaalala..., 2005) viited erinevate tehnoloogiate kasutamisega seotud eetilistele küsimustele puuduvad, siis nagu magistritöö konteksti kirjelduses selgitasin järgib Eesti Rahvusvahelise Sotsiaaltöö Kõrgkoolide Assotsiatsiooni (ingl *International Association of Schools of Social Work*) eestvedamisel loodud dokumendis „Eetikapõhimõtted sotsiaaltöös“ (ingl *Global Social Work Statement of Ethical Principles*, 2018) sätestatut. Suurandmete loomisel, kogumisel, analüüsimisel ja kasutamisel sotsiaaltöö kontekstis on oluline, et nimetatud dokumendi kohaselt kohalduvad kõik dokumendis väljendatud eetikapõhimõtted (st põhimõtted, mis on seotud inimeste vääriskuse austamise, inimõiguste, sotsiaalse õigluse, enesemääramisõiguse ja osalusõiguse edendamisega, konfidentsiaalsuse ja privaatsuse austamisega, inimeste kohtlemisega tervikliku isikuna ning professionaalse usaldusväärsusega) kõigis sotsiaaltöö praktika, hariduse ja uurimise valdkondades, seda nii silmast silma kontaktide puhul kui ka digitehnoloogia või sotsiaalmeedia kasutamisel (Eetikapõhimõtted sotsiaaltöös, 2018).

Sellest tulenevalt on näha, et sotsiaaltöö eetika pakub sotsiaaltöötajatele, nagu Selg (2010: 22) kirjutab, „pidepunkte mõtisklusteks, kahtlustega toimetulekuks ja mitmekülselt kaalutletud otsusteks“ ka suurandmete ja andmetehnoloogiate pakutavate võimaluste kasutamisel sotsiaaltöös. Samas, nagu Reamer (2017) kirjutab, siis kaasnevad tehnoloogia kasutamisega traditsioonilistele sotsiaaltöö eetilistele põhimõtetele (nt privaatsus ja konfidentsiaalsus, teadlik nõusolek, professionaalsus) uued väljakutsed. Sellegipoolest leiab Steiner (2020) näiteks, et vaatamata probleemidele, mis võivad andmetehnoloogiate (Steiner kasutab ingl terminit *digital systems*, aga sisuliselt on tegemist andmetehnoloogiatega) kasutamisega kaasneda (nt automatiseeritud sotsiaalne tõrjutus, algoritmide kallutatus, lugude ja konteksti kadumine, andmekaitse reeglite rikkumine jne) on sotsiaaltöös siiski vähe arutelu valdkonna eetiliste alustalade üle seoses andmetehnoloogiate kasutamisega.

Minu magistritöös analüüsitud artiklitest üksnes pooltes kirjeldatakse suurandmete loomise, kogumise, analüüsimise ja kasutamisega seotud eetilisi probleemkohti sotsiaaltöös ja sedagi enamasti üsna põgusalt. Enim tähelepanu pööratakse privaatsuse (Gillingham ja Graham, 2017; Keddel, 2019; Gillingham 2019b; Kum jt, 2015; Schneider ja Seelmeyer, 2019; Mader, 2015; Zetino ja Mendoza, 2019) ja konfidentsiaalsuse (Gillingham ja Graham, 2017; Keddel, 2019; Kum jt, 2015) ning teadliku nõusoleku andmisega seotud küsimustele (Gillingham ja Graham, 2016; Keddel, 2019; Gillingham, 2019a, 2019b).

Coulton jt (2015), tuginedes Strandburgile (2014) kirjutavad, et kuigi privaatsus ja konfidentsiaalsus on sotsiaaltöös alati olulised olnud, siis suurandmete kontekstis ei piisa andmete kasutamisel teadliku nõusoleku saamiseks tavapärastest meetoditest. Gillingham ja Graham (2017) selgitavad, et teadliku nõusoleku andmine eeldab, et inimene saab aru, kuidas ja mis eesmärgil tema andmeid kasutatakse. Kuna suurandmete kasutamise eesmärgid ja meetodid ei pruugi andmete kogumise hetkel teada olla, siis on teadliku nõusoleku andmine raske (Gillingham ja Graham, 2017). Täiendav väljakutse tuleneb asjaolust, et tihti kasutatakse andmeid, mida on kogutud ühel eesmärgil hoopis teisel eesmärgil, mille jaoks inimene ei ole teadlikku nõusolekut andnud (Keddell, 2015, de Haan ja Connelly, 2014, Dare, 2013; viidatud Gillingham, 2019b). Seda näiteks otsustustoe süsteemide loomisel, mida kasutatakse riskiskooride väljatöötamiseks, et ennetada laste väärkohtlemist (Gillingham, 2019b; Keddell, 2019). Lisaks sellele, et see võib viia erinevate ühiskonnagruppide ebavõrdse kohtlemiseni (Gillingham 2019b; Keddell, 2019) kirjutavad Boyd ja Crawford (2012) erinevate andmestike ühendamisest, et kuna need kõik võivad sisaldada vigu, siis andmestike ühildamisega lihtsalt võimendatakse probleeme veelgi.

Tervikuna, eetiliselt keeruliste küsimuste lahendamisel saavad sotsiaaltöötajad ühest küljest tugineda valdkondlikele eetikakoodeksitele, kuid teisalt ka erinevatele seadustele ja määrustele ning asutuse sisestele suunistele. Nagu Barksy (2017) kirjutab, siis eetikakoodeksid laienevad sotsiaaltööle kogu selle mitmekülgsuses, kuid asutuse siseste juhistega on võimalik sotsiaaltöötajaid täiendavalt toetada.

KOKKUVÕTE

Käesoleva magistritöö eesmärk oli uurida, millised on suurandmete kasutamise võimalused ja suurandmete kasutamisega seotud väljakutsed sotsiaaltöös teadusartiklite põhjal. Uurimustöö läbiviimiseks kasutasin sissejuhatava kirjandusülevaate metoodikat. Selle tulemusel moodustasid uurimuse lõpliku valimi 16 teadusartiklit, millest 10 olid teoreetilised artiklid ja 6 empiirilised uurimused.

Minu magistritöös analüüsitud teadusartiklites käsitletakse sotsiaaltöö valdkonna suurandmetena infosüsteemides ja andmebaasides talletatud andmeid. Teiste andmete (nt interneti rakenduste kasutamisel tekkivate andmete) kasutamisest sotsiaaltöö valdkonnas minu uurimuses analüüsitud artiklites ei räägitud.

Analüüsitud artiklite põhjal saab öelda, et sotsiaaltöö valdkonnas saab suurandmeid kasutada erinevatel eesmärkidel. Mina käsitlesin suurandmete rakendusvõimalusi sotsiaaltöös nelja kategooria lõikes: 1) otsustustoe süsteemid, 2) sotsiaaltöö praktika edendamine, 3) sotsiaaltöö praktika uurimine ja 4) sotsiaalpoliitika kujundamine. Sealjuures kirjeldati analüüsitud artiklites suurandmete potentsiaali just seeläbi, et võimalik on analüüsida suurt kogust andmeid, sh erinevate valdkondade andmeid, mis võimaldab inimeste probleemidest ja vajadustest terviklikuma ülevaate saada ning sotsiaalseid probleeme paremini mõista.

Peamised suurandmete kasutamisega seotud väljakutsed sotsiaaltöös on analüüsitud artiklite põhjal seotud infosüsteemide ja andmebaaside ülesehitusega (st need kas piiravad või liigselt suunavad sotsiaaltöötajate võimalusi klienditööga seonduvat dokumenteerida), sotsiaaltöö valdkonnas loodavate andmete kvaliteediga (sh andmete täpsuse, täielikkuse, esinduslikkuse ja kallutatusega) ning eetiliste küsimustega (sh privaatsuse, konfidentsiaalsuse ja teadlik nõusolekuga).

Sotsiaaltöötajatelt oodatakse, et nad teaksid, millised on suurandmete kasutamise võimalused sotsiaaltöös ning et nad oskaksid suurandmeid analüüsida ja analüüsi tulemusi tõlgendada. Käesolev magistritöö annab seega esimese sissejuhatava ülevaate suurandmete kasutamise võimalustesse ja suurandmete kasutamisega seotud väljakutsetesse sotsiaaltöö valdkonnas.

KASUTATUD KIRJANDUS

Altintas, I. ja Gupta, A. (2020). *Introduction to Big Data*. Loengumaterjal. Kasutatud 01.11.2020, www.coursera.org

Arksey, H., ja O'Malley, L. (2005). Scoping Studies: Towards a Methodological Framework. *International Journal of Social Research Methodology*, 8(1), 19–32.

Aveyard, H. (2014). Why do a literature review in health and social care? *Doing a Literature Review in Health and Social Care, a Practical Guide*, 3rd edition (lk 1–17). Inglismaa: Open University Press.

Barsky, A. E. (2017). Social Work Practice and Technology: Ethical Issues and Policy Responses. *Journal of Technology in Human Services*, 35(1), 8–19.

Baur, N., Graeff, P., Braunisch, L. ja Schweia, M. (2020). The Quality of Big Data. Development, Problems and Possibilities of Use of Process-Generated Data in the Digital Age. *Historical Social Research*, 45(3), 209–243.

Berzin, C. S., Singer, J. ja Chan, C. (2015). Practice Innovation Through Technology in the Digital Age: A Grand Challenge for social Work. *American Academy of Social Work and Social Welfare*. Working Paper No 12. 1–23.

Boyd, D. ja Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>

Carnochan, S., Weissinger, E., Liner-Jigamian, N., Austin, M. J. ja Henry, C. (2019). Using Qualitative Data-Mining to Identify Skillful Practice in Child Welfare Case Records. *Journal of Public Child Welfare*, 13(4), 419–440. <https://doi.org/10.1080/15548732.2018.1509040>

Choo, B.-J., Kim, B.-S., Cha, S.-H., Myung, D.-H. ja Joo, H.-J. (2019). Using multimedia based big data for child abuse prevention system (Based on the establishment of e-child welfare support system). *Multimedia Tools and Applications*, 78(20), 28805–28814.

<https://doi.org/10.1007/s11042-018-6630-0>

Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., Kastner, M. ja Moher, D. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *JOURNAL OF CLINICAL EPIDEMIOLOGY*, 12, 1291–1294.

Coulton, J. C., Goerge, R., Putnam-Hornstein, E. ja de Haan, B. (2015). Harnessing Big Data for Social Good: A Grand Challenge for Social Work. *American Academy of Social Work & Social Welfare, working paper no. 11*, 1–21

Council on Social Work Education. (2018). *2015 Educational Policy and Accreditation Standards*. Kasutatud 20.05.2021, https://www.cswe.org/getattachment/Accreditation/Accreditation-Process/2015-EPAS/2015EPAS_Web_FINAL.pdf.aspx

Crawford, K., Miltner, K. ja Gray, L. M. (2014). Critiquing Big Data: Politics, Ethics, Epistemology. *International Journal of Communication*, 8, 1663–1672.

Daudt, H. M. L., van Mossel, C. ja Scott, S. J. (2013). Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Medical Research Methodology*, 13(1), 1–9.

Devlieghere, J., Roose, R. ja Evans, T. (2020). Managing the electronic turn. *European Journal of Social Work*, 23(5), 767–778. <https://doi.org/10.1080/13691457.2019.1582009>

Eesti Keele Instituut. (2014). Kasutatud 01.11.2020, <http://termin.eki.ee/esterm/concept.php?id=79384&term=big data>

Eetikapõhimõtted sotsiaaltöös. (2018). Kasutatud 20.05.2021, https://www.eswa.ee/wp-content/uploads/2020/06/Eetikapohimotted-sotsiaaltoos_Sotsiaaltoos_4_2020.pdf

Eljas-Taal, K., Veerpalu, A. ja Romanainen, J. (2018). Kas andmetest kujuneb uus varaklass? *Digigeenius*, 29.06.2018. Kasutatud 22.09.2020, <https://digi.geenius.ee/blogi/pikksilm/kas-andmetest-kujuneb-uus-varaklass/>

Gandom, A. ja Heider, M. (2014). Beyond the hype: Big Data concepts, methods and analytics. *International Journal of Information Management*, 35, 137–144.

Gillingham, P. (2019a). Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families? *Child Abuse Review*, 28(2), 114–126. <https://doi.org/10.1002/car.2547>

Gillingham, p. (2011). Decision-making tools and the development of expertise in child protection practitioners: are we ‘just breeding workers who are good at ticking boxes’? *Child and Family Social Work*, 16. 412–421.

Gillingham, P. (2019b). Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge. *Practice*, 31(4), 277–290. <https://doi.org/10.1080/09503153.2019.1575954>

Gillingham, P. (2020). The Development of Algorithmically Based Decision-Making Systems in Children’s Protective Services: Is Administrative Data Good Enough? *British Journal of Social Work*, 50(2), 565–580. <https://doi.org/10.1093/bjsw/bcz157>

Gillingham, P. ja Graham, T. (2017). Big Data in Social Welfare: The Development of a Critical Perspective on Social Work’s Latest “Electronic Turn”. *Australian Social Work*, 70(2), 135–147. <https://doi.org/10.1080/0312407X.2015.1134606>

Goldkind, L. ja Wolf, L. (2015). A Digital Environment Approach: Four Technologies that will Disrupt Social Work Practice. *Social Work*, 60(1), 85–87. <https://doi.org/10.1093/sw/swu045>

Goldkind, L., Wolf, L. ja Freddolino, P. P. (2020). Introduction. L. Goldkind, L. Wolf, P. P. Freddolino (toim), *Digital Social Work* (lk 1–10). Ameerika Ühendriigid: Oxford University Press.

Grant, M. J. ja Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108.

Günther, A. W, Rezazade Mehrizi, H. M., Huysman, M., ja Feldberg, H. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26, 191–209.

Hilbert, M. (2020). *Big Data, Artificial Intelligence, and Ethics*. Loengumaterjal. Kasutatud 09.11.2020, www.coursera.ee

Hou, Zhi-Ping, Chen, Z.-F. ja Luo, Y.-Z. (2020). Applying Big Data Analysis to Discuss the Elderly Social Welfare Service Use Condition and the Relevant Factors. *Revista De Cercetare si Interventie Sociala*, 68, 53–63. <https://doi.org/10.33788/rcis.68.4>

Ishwarappa, K. ja Anuradha, J. (2015). A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. *Procedia Computer Science*, 48, 319–324.

Keddell, E. (2019). Algorithmic justice in child protection: Statistical fairness, social justice and the implications for practice. *Social Sciences*, 8(10), 1–22. <https://doi.org/10.3390/socsci8100281>

Khalil, H., Peters, M., Godfrey, M. C., McInerney, P., Baldini Soares, C. ja Parker, D. (2016). An Evidence-Based Approach to Scoping Reviews. *Worldviews on Evidence-Based Nursing*, 13(2), 118–123.

Kirt, T. (2020). Masinõppe eetodid ja rakendused suurandmetes. A. Masso, K. Tiidenberg ja A. Siibak (toim), *Kuidas mõista andmestunud maailma?* (lk 241–278). Tallinn: TLÜ Kirjastus.

Kitchin, R. (2014). Conceptualising Data. R. Royek (toim), *The Data Revolution. Big Data, Open Data, Data Infrastructures & Their Consequences* (lk 1–26). London: Sage Publications.

Kum, H.-C., Joy Stewart, C., Rose, R. A. ja Duncan, D. F. (2015). Using big data for evidence-based governance in child welfare. *Children and Youth Services Review*, 58, 127–136. <https://doi.org/10.1016/j.childyouth.2015.09.014>

Kukke, A. (2020). *Andres Kukke: see on hädavajalik 21. sajandi kirjaoskus*. Kasutatud, 15.05.2021, <https://www.finantsuudised.ee/arvamused/2020/05/21/andres-kukke-see-on-hadavajalik-21-sajandi-kirjaoskus>

Küingas, P., Haav, M-H., Tepandi, J., Tepandi, L., Lauk, M. ja Tars, T. (2014). *Uuring "Lingitud Eesti"*. Kasutatud 22.09.2020, https://www.riigikantselei.ee/sites/default/files/content-editors/TOF/TOF_uuringud/lingitud_eesti_lopparuanne_2.0.pdf

Laney, D. (2012). *Deja VVVu: Gartner's Original „Volume-Velocity-Variety" Definition of Big Data*. Kasutatud 01.11.2020, <https://community.aiim.org/blogs/doug-laney/2012/08/25/deja-vvvu-gartners-original-volume-velocity-variety-definition-of-big-data>

Lery, B., Haight, J. M. ja Alpert, L. (2016). Four Principles of Big Data Practice for Effective Child Welfare Decision Making. *Journal of Public Child Welfare*, 10(4), 466–474. <https://doi.org/10.1080/15548732.2016.1209149>

Levac, D., Colquhoun, H. ja O'Brien, K. K. (2010). *Implementation Science*, 5(69), 1–9.

McKinsey Global Institute. (2011). *Big Data: the next frontier for innovation, competition, and productivity*. Kasutatud 01.11.2020, <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>

Mader, N. (2015). The Big Data Era and an Integrated Mode of Inquiry for Social Policy-Relevant Research. *I/S: A Journal of Law & Policy for the Information Society*, 11(1), 97–113.

Masso, A., Tiidenberg, K. ja Siibak, A. (2020). Sissejuhatus. A. Masso, K. Tiidenberg ja A. Siibak (toim), *Kuidas mõista andmestunud maailma?* (lk 11–36). Tallinn: TLÜ Kirjastus.

Munn, Z., Peters, D. J. M., Stern, C., Tufanaru, C., McArthur, A. ja Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(143), 1–7. <https://doi.org/10.1186/s12874-018-0611-x>

National Association of Social Workers. (2021). Kasutatud 20.05.2021, <https://www.socialworkers.org/About/Ethics/Code-of-Ethics/Highlighted-Revisions-to-the-Code-of-Ethics>

Naccarato, T. (2010). Child welfare informatics: A proposed subspecialty for social work. *Children and Youth Services Review*, 32(12), 1729–1734.

Nguyen, H. L. (2007). Child Welfare Informatics: A New Definition for an Well Established Practice. *Social Work*, 52(4). 361–313.

O'Brien, K. K., Nayar, A., Colquhoun, H., Wickerson, L., Levac, D., Baxter, L., Tricco, A. C., Straus, S., Moher, D. ja O'Malley, L. (2016). Advancing scoping study methodology: A web-based survey and consultation of perceptions on terminology, definition and methodological steps. *BMC Health Services Research*, 16(305). <https://doi.org/10.1186/s12913-016-1579-z>

Parker-Oliver, D. ja Demiris, G. (2006). Social Work Informatics: A New Specialty. *Social Work*, 51(2). 127–134.

OSKA ülevaade valdkonnaspetsiifiliste IKT-oskuste vajadustest. (2020). Kasutatud 20.05.2021, https://oska.kutsekoda.ee/wp-content/uploads/2020/06/OSKA-%C3%BClevaade-valdkonnaspetsiifiliste-IKT-oskuste-vajadusest_16.06.2020.pdf

Parrott, L. ja Madoc-Jones, I. (2008). Reclaiming information and communication technologies for empowering social work practice. *Journal of Social Work*, 8(2), 181–197. <https://doi.org/10.1177/1468017307084739>

Parton, N. (2008). Changes in the Form of Knowledge in Social Work: From the „Social“ to the „Informational“? *The British Journal of Social Work*, 38(2), 253–269. <https://doi.org/10.1093/bjsw/bcl337>

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D. ja Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence Based Healthcare*, 14, 141–146.

Peters, M. D. J., Godfrey C. M., McInerney, P., Khalil, H., Parker, D., ja Baldini Soares C. (2017).

Guidance for conducting systematic scoping reviews. *International Journal of Evidence Based Healthcare*, 13(3):141-146.

Pham, M. T., Rajic, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A. ja McEwen, S. A. (2014). A Scoping Review of Scoping Reviews: Advancing the Approach and Enhancing the Consistency. *Research Synthesis Methods*, 5, 371–385.

Puusalu, J. (2020). *Suurandmed: olemus ja kasutamise kitsaskohad*. Kasutatud 22.09.2020, <https://digiriul.sisekaitse.ee/bitstream/handle/123456789/2537/2020%2007%20suurandmed-web.pdf?sequence=1&isAllowed=y>

Reamer, F. G. (2005). Documentation in Social Work: Evolving Ethical and Risk-Management Standards. *Social Work*, 50(4). 325–334.

Reamer, F. G. (2017). Evolving Ethical Standards in the Digital Age. *Australian Social Work*, 70(2), 148–159. <https://doi.org/10.1080/0312407X.2016.1146314>

Reamer, F. G. (2013). Social Work in a Digital Age: Ethical and Risk Management Challenges. *Faculty Publications*, 391. 163–172.

Redden, J., Denick, L. ja Warne. H. (2020). Datafied child welfare services: unpacking politics, economics and power. *Policy Studies*, 41(5). 507–526.

Sander, I. (2020). What is Critical Big Data Literacy and How Can It Be Implemented? *Internet Policy Review*, 9(2). 1–22.

Santiago, M. A. ja Smith, J. R. (2019). What Can “Big Data” Methods Offer Human Services Research on Organizations and Communities? *Human Service Organizations: Management, Leadership & Governance*, 34(4), 344–356. doi:10.1080/23303131.2019.1674756

Schneider, D. ja Seelmeyer, U. (2019). Challenges in Using Big Data to Develop Decision Support Systems for Social Work in Germany. *Journal of Technology in Human Services*,

37(2–3), 113–128. <https://doi.org/10.1080/15228835.2019.1614513>

Schwartz, I. M., York, P., Nowakowski-Sims, E. ja Ramos-Hernandez, A. (2017). Predictive and prescriptive analytics, machine learning and child welfare risk assessment: The Broward County experience. *Children and Youth Services Review*, 81, 309–320.

Shafer, T. (2017). *The 42 V's of Big Data and Data Science*. Kasutatud 01.11.2020, <https://www.elderresearch.com/blog/42-v-of-big-data>

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.

Sotsiaalalatoötaja eetikakoodeks. (2015). Kasutatud 20.05.2021, <https://www.eswa.ee/wp-content/uploads/2016/11/lisa-3sotsiaalala-tootaja-eetikakoodekspdf.pdf>

Sutton, A., Clowes, M., Preston, L. ja Booth, A. (2019). Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Information and Libraries Journal*, 36, 202–222.

Standards for Technology in Social Work Practice
https://www.socialworkers.org/includes/newIncludes/homepage/PRA-BRO-33617.TechStandards_FINAL_POSTING.pdf

Tartu Ülikooli raamatukogu koolitusmaterjalid ja juhendid (i.a). Kasutatud 20.05.2021, <https://utlib.ut.ee/koolitusmaterjalid-ja-juhendid>

Taylor, S., Battis, C., Carnochan, S., Balk, M., Austin, M. J. ja Henry, C. (2019). Exploring trauma-informed practice in public child welfare through qualitative data-mining of case records. *Journal of Public Child Welfare*, 13(3), 325–344. <https://doi.org/10.1080/15548732.2018.1500967>

Teare, G. ja Taks, M. (2020). Extending the Scoping Review Framework: A Guide for Interdisciplinary Researchers. *International Journal of Social Research Methodology*, 23(3), 311–315.

The Digital Competence Framework. (i.a). Kasutatud 23.05.2021,
file:///C:/Users/pisike/Downloads/KJNA28558ENN.en.pdf

Tricco, C. A., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Kastner, M., Levac, D., Ng, C., Pearson Sharpe, J., Wilson, K., Kenny M., Warren, R., Wilson, C., Stelfox, T. H. ja Straus, E. S. (2016). A scoping review on the conduct and reporting of scoping reviews. *BMC Medical Research Methodology*, 16(15), 1–10. doi10.1186/s12874-016-0116-4

Xiao, Y. ja Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *JOURNAL OF PLANNING EDUCATION AND RESEARCH*, 39(1), 93–112.
<https://doi.org/10.1177/0739456X17723971>

Zetino, J. ja Mendoza, N. (2019). Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare. *Social Work in Public Health*, 34(5), 409–417.

Zhi-Ping HOU, Zhi-Fei CHEN, ja Yu-Zhou LUO. (2020). Applying Big Data Analysis to Discuss the Elderly Social Welfare Service Use Condition and the Relevant Factors. *Revista de Cercetare si Interventie Sociala*, 68, 53–63.

5 LISAD

5.1 LISA 1. Andmete kaardistamise vorm

	Teadusartikli pealkiri	Autor	Aasta	Ajakiri	Riik, kus uurimus avaldati või kus uurimus läbi viidi	Andmete liik (uurimuses kasutati või uurimus räägib kasutamisest)	Sotsiaaltöö valdkond	Kas artiklis käsitletakse suurandmete kasutamisega seotud eetilisi küsimusi sotsiaaltöös?	Kas artiklis tuuakse näiteid selle kohta, kuidas on suurandmeid sotsiaaltöö valdkonnas kasutatud või kuidas oleks võimalik suurandmeid kasutada?
1	“Algorithmic Justice in Child Protection: Statistical Fairness, Social Justice and the Implications for Practice”	Emily Keddell	2019	Social Sciences	Uus-Meremaa	Administratiivsed andmed	Lastekaitse valdkond	Jah	Jah (otsustustoe süsteemid).
2	“Health social workers’ assessments as part of a specialized pain rehabilitation: a clinical data mining study”	Kjerstin Larsson, Ruth Kusec Fredriksson ja Kerstin Sjogren Fugl-Meyer	2019	Social Work in Health Care	Rootsi	Administratiivsed andmed	Sotsiaaltöö tervishoius	Ei	Jah (empiiriline uurimus).
3	“Predictive and prescriptive analytics, machine learning and	Ira M. Schwartz, Peter York,	2017	Children and Youth	Ameerika Ühendriigid	Administratiivsed andmed	Lastekaitse valdkond	Ei	Jah (empiiriline uurimus).

	child welfare risk assessment: The Broward County experience”	Eva Nowakowski-Sims ja Ana Ramos-Hernandez		Services Review					
4	“Decision Support Systems, Social Justice and Algorithmic Accountability in Social Work: A New Challenge”	Philip Gillingham	2019	Practice	Austraalia	Administratiivsed andmed	Sotsiaaltöö valdkond tervikuna	Jah	Jah (otsustustoe süsteemid).
5	“The Development of Algorithmically Based Decision-Making Systems in Children’s Protective Services: Is Administrative Data Good Enough?”	Philip Gillingham	2020	British Journal of Social Work	Austraalia	Administratiivsed andmed	Lastekaitse valdkond	Ei	Jah (otsustustoe süsteemid).
6	“Exploring trauma-informed practice in public child welfare through qualitative data-mining of case records”	Sarah Taylor, Claire Battis, Sarah Carnochan, Colleen Henry, Margaret Balk ja Michael J. Austin	2019	Journal of Public Child Welfare	Ameerika Ühendriigid	Administratiivsed andmed	Lastekaitse valdkond	Ei	Jah (empiiriline uurimus).
7	“Using Qualitative Data-Mining to Identify Skillful	Sarah Carnochan, Erika	2019	Journal of Public	Ameerika Ühendriigid	Administratiivsed andmed	Lastekaitse valdkond	Ei	Jah (empiiriline uurimus).

	Practice in Child Welfare Case Records”	Weissinger , Colleen Henry, Nicole Liner-Jigamian ja Michael J. Austin		Child Welfare					
8	“Using multimedia based big data for child abuse prevention system (Based on the establishment of e-child welfare support system)”	Bong -Jo Choo, Hae-Jong Joo, Bong-Soo Kim, Si-hwan Cha ja Do-Hyun Myung	2019	Multimedia Tools and Applications	Korea	Administratiivsed andmed	Lastekaitse valdkond	Ei	Jah (empiiriline uurimus).
9	“Using big data for evidence-based governance in child welfare”	Hye-Chung Kum, C. Joy Stewart, Roderick A. Rose ja Dean F. Duncan	2015	Children and Youth Services Review	Ameerika Ühendriigid	Administratiivsed andmed	Lastekaitse valdkond	Jah	Jah (kirjeldatakse suurandmete kasutamise võimalusi).
10	“Four Principles of Big Data Practice for Effective Child Welfare Decision Making”	Bridgette Lery, Jennifer M. Haight ja Lily Alpert	2016	Journal of Public Child Welfare	Ameerika Ühendriigid	Administratiivsed andmed	Lastekaitse valdkond	Ei	Ei
11	“Challenges in Using Big Data to Develop Decision Support Systems for Social Work in Germany”	Diana Schneider ja Udo Seelmeyer	2019	Journal of Technology in Human Services	Saksamaa	Administratiivsed andmed	Sotsiaaltöö puuetega inimestega	Jah	Jah (otsustustoe süsteemid).

12	“The Big Data Era and an Integrated Mode of Inquiry for Social Policy-Relevant Research”	Nicholas Mader	2015	I/S: A Journal of Law & Policy for the Information Society	Ameerika Ühendriigid	Administratiivsed andmed	Sotsiaaltöö valdkond tervikuna	Jah	Jah (kirjeldatakse suurandmete kasutamise võimalusi).
13	“Big Data in Social Welfare: The Development of a Critical Perspective on Social Work's Latest “Electronic Turn”	Philip Gillingham ja Timothy Graham	2016	Australian Social Work	Austraalia	Administratiivsed andmed	Sotsiaaltöö valdkond tervikuna	Jah	Jah (otsustustoe süsteemid, kirjeldatakse suurandmete kasutamise võimalusi).
14	“Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare”	Jeffrey Zetino & Natasha Mendoza	2019	Social Work in Public Health	Ameerika Ühendriigid	Administratiivsed andmed	Sotsiaaltöö valdkond tervikuna	Jah	Jah (kirjeldatakse suurandmete kasutamise võimalusi).
15	Applying Big Data Analysis to Discuss the Elderly Social Welfare Service Use Condition and the relevant factors	Zhi-Ping HOU, Zhi-Fei CHEN, Yu-Zhou LUO	2020	Revista de Cercetare si Interventie Sociala	Hiina	Administratiivsed andmed ning valitsusasutuste analüüsid ja uuringud	Sotsiaaltöö eakatega	Ei	Jah (empiiriline uurimus).
16	Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families	Philip Gillingham	2019	Child Abuse Review	Austraalia	Administratiivsed andmed	Lastekaitse valdkond	Jah	Jah (otsustustoe süsteemid).

Mina, Katrin Siider, annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Suurandmete kasutamise võimalused sotsiaaltöös: sissejuhatav kirjandusülevaade“, mille juhendaja on Merle Linno, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

1. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
2. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Katrin Siider

25.05.2021

