

Tiedonhaku aidoissa työtehtävissä: Kenttätutkimus tehtävätyypin ja monimutkaisuuden vaikutuksista

Miamaria Saastamoinen

Tampereen yliopisto

miamaria.saastamoinen@helsinki.fi

<http://orcid.org/0000-0002-7371-1490>

FM Miamaria Saastamoisen informaatiotutkimuksen ja interaktiivisen median alaan kuuluva väitöskirja Information searching in authentic work tasks: A field study on the effects of task type and complexity (Tiedonhaku aidoissa työtehtävissä: Kenttätutkimus tehtävätyypin ja monimutkaisuuden vaikutuksista) tarkastettiin 14.1.2017 Tampereen yliopistossa. Vastaväittäjänä toimi professori Katriina Byström (Oslo and Akershus University College of Applied Sciences, Norway) ja kustoksena yliopistolehtori Jaana Kekäläinen (Tampereen yliopisto). Väitöskirja on julkaistu sarjassa Acta Universitatis Tamperensis ja se on luettavissa myös Tampereen yliopiston Tampub-julkaisuarkistossa osoitteessa <http://urn.fi/URN:ISBN:978-952-03-0310-5>.

Asiasanat: tiedonhaku; tiedonhankinta; tietotyö; kenttätutkimus

Arvoisa vastaväittäjä, kustos ja yleisö; arvoisat tiedon ja tieteen ystävät. Tässä esityksessä käsittelen lyhyesti väitöskirjani taustaa, tutkimuskysymyksiä ja tutkimusmenetelmiä. Väitöskirja kuuluu informaatiotutkimuksen alaan. Alan nimi, informaatiotutkimus, on kuvaava: alalla tutkitaan informaatiota, tietoa, ja erityisesti tiedon löydettävyyden kysymyksiä, joten ala on esimerkiksi arkistotoiminnan, kirjastojen ja tiedonhakujärjestelmien ydinkysymysten äärellä.

Artikkeli on lisensoitu Creative Commons Nimeä-EiKaupallinen-JaaSamoin 4.0 Kansainvälinen -lisenssillä

Pysyvä osoite: <http://dx.doi.org/10.23978/inf.63193>

Väitöstutkimukseni kuuluu informaatiotutkimuksen alan sisällä tiedonhankintatutkimuksen ja tiedonhakututkimuksen piiriin. Yleiskielen termeinä tiedonhankintaa ja tiedonhakua käytetään usein synonyymeina, mutta informaatiotutkimuksessa ne viittaavat hieman eri asioihin, erityisesti siten, että tiedonhaku viittaa sähköisten tietojärjestelmien käyttöön, ja tiedonhankinta muunlaisten tiedonlähteiden käyttöön. Perinteinen tiedonhankintatutkimus on yhteiskuntatieteellistä tutkimusta, esimerkiksi kirjastonkäyttötutkimusta. Tiedonhankintatutkimuksen piirissä on analysoitu, millaisia tietoja ihmiset haluavat, kuinka he niitä hankkivat ja kuinka he lopulta tiedosta hyötyvät eli kuinka he tietoa käyttävät. Perinteinen tiedonhakututkimus taas keskittyy tietokoneen kautta käytettäviin tiedonhaku- ja tallennusjärjestelmiin ja niiden hyvinkin tekniseen kehittämiseen laboratorio-olosuhteissa. Kun tiedonhakujärjestelmän käyttäjä tuodaan tutkimukseen mukaan, on kyse vuorovaikutteisesta tiedonhausta. Tyypillisesti käyttäjälle annetaan hakutehtävä eli kuvaus tilanteesta, johon hänen tarvitsee hakea tietoa. Hän hakee tietoa järjestelmästä hakusanoilla eli hakulausekkeilla ja arvioi löydettyjen dokumenttien hyödyllisyyttä tilanteeseen nähden. Tällä tavalla katsottuna tiedonhankinta- ja tiedonhakututkimus ovat aika kaukana toisistaan. Todellisissa tilanteissa oikeat tiedontarvisijat eivät kuitenkaan erittele tiedon etsimisen vaiheita, vaan he kokevat tiedontarpeen, toimivat sen tyydyttämiseksi jollain lailla jos toimivat, esimerkiksi käyttävät sähköisiä tai muita lähteitä, ja jatkavat sitä mitä ikinä olivatkaan tekemässä. Jokseenkin eksaktin tutkimuksen tarpeisiin on toki tärkeää eritellä tätä prosessia yksityiskohdaisemmin. Mielestäni on kuitenkin hyvä pitää mielessä myös kokonaiskuva eli se todellisuus, jota varten koko tutkimusala on olemassa.

Tutkimukseni tavoitteena on ollut kuvailla ja ymmärtää osaltaan tätä kokonaiskuva. Tutkimuksessani analysoin, miten työtehtävien piirteet vaikuttavat tiedonhankintaan ja tiedonhakuun. Henkilökohtainen kiinnostuksen kohteeni on erityisesti sähköisissä tiedonhakujärjestelmissä ja siinä, miten niitä käytetään. Kuten väitöskirjaani piirtämässäni kuvassa esitän, tehtävälähtöisen tiedonhaun tutkimuksessa tiedonhakuun liittyy kuitenkin myös muunlainen tiedonhankinta ja tietenkin työtehtävät. Teoreettinen pohja, jolta työni ponnistaa, on yhdistelmä aiemman kirjallisuuden ehdottamista malleista, joiden peruseriaatteista seuraavaksi muutama sana. Tietotyössä tiedon käsittelemisellä on keskeinen rooli. Tietotyö koostuu työtehtävistä, joilla on kullakin omat piirteensä. Väitöskirjassa käsittelen erityisesti työtehtävän monimutkaisuuden vaikutusta tiedonhakuun. Minun aineistossani työtehtävällä tarkoitetaan tutkittavan itse nimeämää eli määrittelemää kokonaisuutta. En siis ole pyytänyt tutkittavaa vaikkapa tekemään matkalaskua, vaan jos hänellä on suoritettavanaan matkalasku aineistonkeruun aikana, hän kirjaa tiedon vapaa-muotoisesti tehtävälomakkeelle. Työtehtävät luonnollisesti suoritetaan käyttämällä erilaisia välineitä erilaisiin tarkoituksiin. Työtehtävässä pitää esimerkiksi kirjoittaa, laskea tai keskustella kollegan kanssa. Joskus työtehtävässä törmätään tiedontarpee-

seen, eli tehtävässä etenemiseksi tarvitaan tietoa. Silloin tiedonhankkija eli tutkittava saattaa ruveta hankkimaan tietoa eli käyttämään tiedonlähteitä tiedontarpeen edellyttämällä tavalla. Tiedonhankinnassa saattaa tulla tilanteita, joissa hakulausekkeen tekeminen sähköiseen tietojärjestelmään nähdään parhaaksi vaihtoehdoksi, jolloin siirrytään tiedonhakututkimuksen kiinnostuksen alueelle.

Tehtävälähtöistä tiedonhakua ja tiedonhankintaa on kyllä tutkittu empiirisesti aikaisemminkin, joten on hyvä tässä yhteydessä kertoa aiemmasta tutkimuksesta, jotta on mahdollista ymmärtää, miten tutkimukseni poikkeaa siitä. Kuten mainitsin, kontrolloiduissa olosuhteissa voidaan tehdä tiedonhakukokeita siten, että osallistujina ovat oikeat ihmiset, jotka tosin useimmiten suorittavat keinotekoisia haktehtäviä, joissa voi olla enemmän tai vähemmän realistinen taustakuvaus hakua helpottamassa. Toinen tyypillinen tiedonhaun tutkimustapa ovat lokianalyysit. Lohki on siis automaattisesti kerättyä päiväkirjaa yksittäisen tietokoneen kaikista tapahtumista tai usean tietokoneen tapahtumista yhdessä tietojärjestelmässä riippuen siitä, mihin lokia keräävä ohjelma on asennettu. Voidaan esimerkiksi analysoida, millaisia hakulausekkeitä ihmiset tekevät internetin hakukoneisiin. Tällöin nähdään kuitenkin vain yhden järjestelmän, eli kyseisen hakukoneen käyttö, eikä tyypillisesti voida tietää, millaisessa tilanteessa hakija oli haun tehdessään, eli miksi hän haun teki, mihin suurempaan kokonaisuuteen se liittyi. Tätä kontekstia valottaakseen jotkin tutkimukset käyttävät aineistonkeruumenetelmiä, joissa tutkittavat kertovat, mitä he tietoa hankkiessaan tekevät ja miksi. Heitä voidaan haastatella, he saattavat täyttää päiväkirjaa tai kyselylomakkeita. Tällaiset menetelmät tarjoavat hyvin tietoa siitä, miksi ihmiset tietoa hakevat, mutta itse asiassa tutkija ei voi tietää, ovatko tutkittavat tehneet juuri niin kuin he kertovat tehneensä. Ihmisen muistin valikointivuus on tässä keskeinen tekijä. Tiedonhankintaan ei välttämättä kiinnitä huomiota, eikä siksi muista siitä enää yksityiskohtia edes muutamaa tuntia myöhemmin. Muistatko sinä viimeisimmän Google-hakusi? Entä sitä edellisen? Montaako hakulauseketta eli hakusanojen yhdistelmää käytit internetissä viime tiistaina? Omassa tutkimuksessani huomasin myös, että ihmiset haluavat selittää toimintansa mahdollisimman järkevästi jälkikäteen, jolloin esimerkiksi ajatusvirheet tai muut ongelmat jäävät tutkijalta pimentoon. Jotkut tutkijat taas ovat tuoneet tutkittavansa sinänsä keinotekoisiihin ja kontrolloituihin hakutilanteisiin, joissa tutkittavat kuitenkin saavat hakea tietoa aitoihin tehtäviinsä, tyypillisesti opiskelutehtäviin, liittyen. Tällaisessa asetelmassa voidaan siis todentaa tapahtunut, niin sanotusti tosielämän, tiedonhaku. Tässäkin lähestymistavassa on kuitenkin ongelmansa: tiedonhaku tapahtuu määrättyinä hetkenä, ei silloin, kun olisi luonnollisesti sen aika, ja käytössä ovat vain ne tiedonlähteet, jotka tutkijalla on sillä hetkellä tutkittaville tarjota, vaikkapa vain yksi tietokanta.

Jos siis haluamme ymmärtää, miksi ja milloin tiedonhakua tosielämässä tapahtuu ja millaista se on, mitä pitäisi tehdä? Väitöskirjassa esitän käyväksi ratkaisuksi

kenttätutkimusmenetelmää. Kenttätutkimuksella tarkoitetaan sitä, että tutkija jalkautuu tutkittaviensa joukkoon seuraamaan heidän tekemisiään, tässä tapauksessa heidän työtään. Varsinkin laaja-alaiset kenttätutkimukset ovat suhteellisen harvinaisia tiedonhankinta- ja tiedonhakututkimuksen piirissä, ehkäpä siksi että ne vaativat erityisen paljon aikaa: tiedonhakutilannetta ei keinotekoisesti synnytetä vaan sellaisen ilmaantumista on vain odotettava. Koska jokaisessa yksittäisessä aineistonkeruumenetelmässä, kuten haastatteluissa, kyselyissä tai lokin keräämisessä on puutteensa, olen hyödyntänyt useita menetelmiä samanaikaisesti mahdollisimman todenmukaisen kuvan saamiseksi. Olen istunut tutkittavien vieressä havainnoimassa omin silmin heidän työtään, tehnyt muistiinpanoja ja nauhoittanut heidän kertomaansa, he ovat kantaneet kaulassaan kameraa, joka kertoo missä he liikkuvat, heidän tietokoneelleen on asennettu tietokoneen tapahtumista tietoa keräävä loki-ohjelma, heidän näyttöään on videoitu, heitä on haastateltu, ja he ovat kertoneet sähköisillä kyselylomakkeilla työtehtävistään.

Tutkimukseni muodostuu kahdesta erillisestä aineistosta. Ensimmäinen, pienempi aineisto, kerättiin kaupungin hallinnossa, jossa lokien keräämiseen ei ollut lupaa. Tällöin tärkeimmät aineistonkeruumenetelmät olivat suora havainnointi ja tehtäväkohtaiset kyselylomakkeet. Toinen aineisto on määrällisesti moninkertainen ensimmäiseen aineistoon nähden. Se kerättiin useissa eri organisaatioissa sekä julkisella että yksityisellä sektorilla. Näissä organisaatioissa lokin kerääminen oli sallittua, joten loki ja kyselyvastaukset ovat toisessa aineistossa tärkeintä aineistoa analyysin kannalta. Kuitenkin molemmissa aineistoissa kaikki erilaiset aineistotyyppit tukivat toisiaan erityisesti alustavassa analyysissa.

Aineistonkeruumenetelmiä on suhteellisen helppoa kuvailla. Sen sijaan aineiston analyysia on vaikeampi lyhyesti selittää. Kuvitelkaa, että teitä seurattaisiin vaikkapa päivän ajan. Kuinka helppoa ulkopuolisen olisi jäsentää, mitä teitte ja miksi? Aluksi analyysissa kyse oli siis käytännössä salapoliisityöstä. Minun tuli tutkimuksen tavoitteiden mukaisesti määritellä, mitä ilmiöitä aineistosta voi analysoida, mitä sieltä on siis löydettävissä. Tällainen ilmiö voisi olla vaikkapa hakutehtävä. Se on määriteltävä aineistoon sopivasti. Aiemmassa esimerkissä kuvailemani määritelmä hakutehtävästä oli yksinkertaisesti tutkijan valmiiksi tekemä tilannekuvaus. Tällainen määritelmä ei tietenkään olisi sopinut aineistolleni, koska tässä tutkimuksessa ei ollut keinotekoisia tutkijan väliintuloja. Kun uusi määritelmä oli luotu, piti aineistosta etsiä määritelmän mukaiset tapahtumat ja luokitella ne, jotta päästiin määrälliseen analyysiin. Vaikka aineistoni on lähtökohtaisesti nimenomaan laadullista, pidin tärkeänä sitä, että voin jalostaa aineistoa niin, että määrällinen analyysi on mahdollista. Tämä helpottaa tulosten yleistämistä ja vertailua muihin tutkimuksiin. Tyypillinen laadullinen analyysihan usein laajasti kuvailee tutkittavia ilmiöitä ja höyryttää niitä esimerkeillä. Minä pyrin tästä askelta abstraktimmalle tasolle eli siihen, että aineistoa voi sanallisen kuvailun lisäksi kuvailla numeroin. Jos aineistolla

pääsee laskemaan, voi mahdollisesti hyödyntää tilastomenetelmiä. Tilastomenetelmät auttavat jäsentämään ja tulkitsemaan löydöksiä. Tilastotesteillä voidaan tutkia, johtuvatko saadut luvut ja lukujen erot sattumasta. Sattuma ei useimmiten ole toivottu tulos, vaan tutkija haluaa löytää niin kutsutun tilastollisesti merkitsevän tuloksen. Minäkin olen käyttänyt tilastollisia menetelmiä analyysin loppuvaiheessa aina tarvittaessa.

Väitöskirjani koostuu konkreettisesti kuudesta artikkelista, jotka on julkaistu tieteellisissä, vertaisarvioituissa julkaisuissa. Esittelen seuraavaksi lyhyesti nämä artikkelit. Kolme ensimmäistä artikkelia on kirjoitettu ensimmäisestä aineistosta, kolme seuraavaa toisesta aineistosta.

Ensimmäinen artikkeli on läpileikkaus ensimmäisestä aineistosta. Siinä analysoimme, miten työtehtävien monimutkaisuus vaikuttaa käytettyihin tiedonlähteisiin, tehtyihin hakulausekkeisiin ja työtehtävässä kohdattuihin ongelmiin. Analysoitava aineisto oli havainnointimuistiinpanoja kentältä. Ensimmäinen artikkeli on väitöskirjan ainoa konferenssijulkaisu, muut on julkaistu tieteellisissä lehdissä. Artikkeli oli myös ehdolla konferenssin parhaaksi artikkeliksi.

Toisessa artikkelissa analysoimme, mitä tietoa tutkittavat tarvitsivat työtehtäviensä suorittamiseksi. Tiedontarpeitaan he kuvailivat sähköisillä, tehtäväkohtaisilla kyselylomakkeilla vapaamuotoisesti. Näistä kuvailuista tein tiedontarveluokituksen, ja laskin, miten tiedontarpeet vaihtelivat tehtävän monimutkaisuuden muuttuessa. Analysoin sekä ennen tehtävää kuvailtuja eli oletettuja tiedontarpeita että tehtävän jälkeen kuvailtuja tiedontarpeita sekä näiden kahden eroja.

Kolmas artikkeli käsittelee niin ikään tehtäväkohtaisia kyselylomakevastauksia, tällä kertaa siitä näkökulmasta, miten tiedonlähteiden käyttö vaihtelee tehtävän monimutkaisuuden myötä. Artikkeli sisältää myös aiemman tutkimuksen perusteella tehdyn kartoituksen siitä, millaisia kirjallisuudessa tarjotut tiedonlähdeluokitukset ovat aiemmin olleet.

Neljännessä artikkelissa siirrytään toiseen, isompaan ja monipuolisempaan aineistoon. Tehtävän monimutkaisuuden lisäksi toisesta aineistosta nousi esiin tehtävätyyppiluokitus, jota käytämme riippumattomana eli selittävänä tekijänä monimutkaisuuden ohella. Artikkelissa analysoimme työtehtävissä esiintyviä hakutehtäviä. Hakutehtävällä tarkoitan siis hakulausekkeiden joukkoa, joka tehdään samaan tavoitteeseen pyrkien, eli taustalla on jokin yhtenäinen tiedontarve. Hakutehtävien ominaisuuksista artikkelissa käsitellään käytettyjen tiedonlähteiden tyyppiä ja määrää, tiedontarvetta hakutehtävän taustalla sekä hakutehtävän suorittamisprosessin tyyppiä.

Viideskin artikkeli käsittelee hakutehtäviä, kuitenkin hakusanojen näkökulmasta. Artikkelissa analysoimme hakulausekkeiden lukumäärää, tyyppiä ja pituutta. Lisäksi analysoimme hakutehtävien lukumäärää per työtehtävä. Tämän tyyppiset asiat ovat usein kiinnostuksen kohteena perinteisessä tiedonhaku tutkimuksessa. Nämä

muuttujat ovat hyvin suoraviivaisesti määriteltävissä hakukoneen lokia analysoidessa tai laboratorio-olosuhteissa, mutta niiden soveltaminen rikkaaseen tosielämän aineistoon olikin haastavampaa kuin saattaisi kuvitella.

Kuudes artikkeli käsittelee ilmiötä, jota englanniksi kutsutaan nimellä *dwell time*. Se voitaisiin kääntää esimerkiksi pysähdysajaksi tai viipymisajaksi. Se on suosittu muuttuja perinteisessä tiedonhakututkimuksessa. Perinteinen lähestymistapa on analysoida, vaikuttaako löydetyn dokumentin, esimerkiksi webbisivun, koettu hyödyllisyys sen lukemisaikaan, siis siihen, kauanko dokumentissa viivytään. Me laajensimme tätä käsitystä siten, että analysoimme, kauanko tutkittavat käyttivät kaikkia tiedonlähteitään, ei vain siis yhtä ja yhteen tarkoitukseen, ja miten tähän olivat yhteydessä työtehtävän tyyppi ja työtehtävän monimutkaisuus.

Väitöskirjan yhteenveto-osassa käsitelän kootusti yhtä aikaa kaikkia kuutta artikkelia. Artikkeleiden avulla vastaan neljään tutkimuskysymykseen. Ensimmäisen tutkimuskysymyksen puitteissa analysoin yleisellä tasolla aineiston ominaisuuksia: millaisia työtehtävän piirteitä aineistosta löytyy, paljonko tutkittavat käyttävät tiedonlähteitä, paljonko he hakevat tietoa ja millaisia ongelmia he työssään kohtaavat. Toisessa tutkimuskysymyksessä käsitelän tiedontarvetyyppejä, kolmannessa taas sitä, miten tutkittavat reagoivat tiedontarpeisiinsa, eli miten he tietoa hankkivat ja hakevat. Viimeinen tutkimuskysymys keskittyy aineistossa ilmeneviin hakusanoihin. Ylipäänsä analysoin tiedontarpeita, tiedonlähteitä ja tiedonhakua siitä näkökulmasta, miten työtehtävän piirteet niihin vaikuttavat.

Tutkimukseni keskittyy siis kuvaamaan empiirisiä löydöksiä. Eri tutkimusten tuloksia vertailemalla voidaan rakentaa malleja tiedonhakukäyttäytymisestä ja edelleen teorioita siitä. Lisäksi, koska vastaavia tutkimuksia on vähän, täytyi tutkimusta varten kehittää ainutlaatuinen menetelmäpaketti. Toivon, että työstäni on hyötyä tulevaisuuden tiedonhaun tutkimukselle vähintään niin, että kokemukseni kenttätyöstä antaisivat innoitusta uusille tutkimuksille. Tiedonhaun tai laajemmin informaatiotutkimuksen ala on todella tärkeä nykyiselle informaatioyhteiskunnalle. Parhaimmillaan esimerkiksi tiedonhakujärjestelmät ovat silloin, kun ne ovat näkymättömiä, kun tieto noin vain löytyy, eikä tavallisen tiedonhakijan tarvitse edes ymmärtää, miten se tapahtui. Tämän onnistuneen kokemuksen takana ovat tutkimus ja tiedonhaun ammattilaiset.

Lopuksi en voi olla mainitsematta, että tänään on merkittävän suomalaisen artistin, Rauli Badding Somerjoen kuoleman vuosipäivä. Erityisesti Somerjoen uran varhaisvaiheen laulut ovat vahvasti kantaaottavia. Moni ei ehkä tiedäkään, että hän on ottanut kantaa myös tieteeseen. Eräs laulu vuodelta 1969 kertoo, että tieteellisen teoksen kohdalla pitää kysyä: Sisältääkö se jotain suuretta tai lukua koskevaa loogillista ajattelua? Sisältääkö se jotain tosiseikkoja ja olemassaoloa koskevaa kokemusajattelua? Jos vastaus on ei, teos pitää laulun mukaan heittää tuleen ”sillä se ei voi sisältää muuta kuin viisastelua ja harhaluuloa”.

Pyydän Teitä, arvoisa professori Katriina Byström, Informaatiotieteiden yksikön, nykyisen Viestintätieteiden tiedekunnan, määräämänä vastaväittäjänä esittämään ne muistutukset, joihin katsotte väitöskirjani antavan aiheita.