

Johanna Lehtimäki

# TYÖPERÄISTEN SAIRAUKSIEN REKISTERIN KÄYTÖN ANALYYSI TIETOPALVELUPYYNTÖJEN PERUSTEELLA

Informaatioteknologian ja viestinnän tiedekunta  
Pro gradu -tutkielma  
Huhtikuu 2021

# TIIVISTELMÄ

Johanna Lehtimäki: Työperäisten sairauksien rekisterin käytön analyysi tietopalvelupyyntöjen perusteella  
Pro gradu -tutkielma  
Tampereen yliopisto  
Informaatiotutkimuksen maisteriohjelma  
Huhtikuu 2021

---

Työperäisten sairauksien rekisteri eli TPSR on Työterveyslaitoksen ylläpitämä rekisteri, joka sisältää tietoa ammattitaudeista ja ammattitautiepäilyistä. Rekisterin pitäminen perustuu lakiin ja sen tietoja käytetään muun muassa työterveysalan tutkimuksessa. Rekisteristä julkaistaan vuosittain virallinen tilasto ammattitaudeista ja ammattitautiepäilyistä, annetaan tietoja EU:lle sekä tehdään tarkempia tilastoja tietopalvelupyyntöjen perusteella. Tutkimuksen tavoitteena oli selvittää rekisterin käyttöä, jotta jatkossa voitaisiin kehittää rekisterin julkaisuja.

Aikaisempi tutkimus on keskittynyt arkistojen ja asiakirjojen käyttäjiin, eikä vastaavanlaisia rekisterin käyttöön liittyviä tutkimuksia löydy. Tutkimuskysymykset koskivat rekisterin käyttäjiä eli tiedon pyytäjien taustaa sekä pyyntöjen sisältöä. Tutkimuksessa analysoitiin neljän vuoden aikana tulleiden tietopalvelupyyntöjen sisällöt, eli selvitettiin mihin tarkoitukseen tietoja pyydetään, kuka niitä pyytää ja kuinka paljon.

Tutkimus oli kartoittava kvantitatiivinen analyysi. Alun perin sanallisesta aineistosta muodostettiin havaintomatriisi niin, että kaikki havainnot muunnettiin ensin numero- ja merkkimuotoisiksi. Näin myös tulokset voitiin esittää numeerisessa muodossa. Tietopalvelupyynnöt palasteltiin omiksi tiedoikseen käyttäjän ja pyynnön sisällön mukaan sekä luokiteltiin. Pyyntöjen sisällöt muokattiin sammanlaisiksi muuttujiksi, kuin mitä rekisteri sisältää.

Tuloksista selvisi, että pyyntöjä tulee vuosittain melko tasaisesti. Tietopalvelupyyntöjä tekivät eniten työssäkäyvät ja vähemmän tietoja pyysivät opiskelijat ja erikoistuvat lääkärit. Pyyntöjen tarkoitus liittyi useimmiten asiantuntijakäyttöön tai tutkimushankkeeseen, ja tähän ryhmään kuuluivat myös median tekemät pyynnöt. Tietoja pyydetään myös opinnäytetöitä ja erilaisia (painettuja) julkaisuja varten. Suosituimmat vuosirajaukset pyynnöissä olivat 10, 5 ja 1 eli viimeisin vuosi.

Tulosten mukaan suosituin hakuheito oli sairaus ja eniten tietoja haluttiin ihotaudeista, infektio-  
taudeista sekä astmoista. Altisteen mukaan pyydettiin tietoja toiseksi useimmin, ja yli puolet näistä koski kemiallisten altisteiden ryhmään kuuluvia altisteita. Kemiallisiin altisteisiin kuuluvat esimerkiksi metallit, asbesti, liimat, pesuaineet ja kasviperäiset altisteet. Myös biologisten altisteiden mukaan pyydettiin tietoja, esimerkiksi homeiden, virusten ja alkueläimien aiheuttamia tapauksia. Tilastoja pyydettiin lisäksi toimialojen ja ammattien mukaan, mutta niistä ei noussut esiin selkeitä suosikkeja. Opiskelijoita kiinnostivat tilastot eniten opiskelemansa ammatin mukaan, kun taas erikoistuvia lääkäreitä kiinnostivat eniten tilastot sairauden tai altisteen mukaan haettuna.

Tutkimus vahvisti, että vuosijulkaisuissa kannattaa edelleen esitellä uusimpia tuloksia sairauksien myötä sekä ottaa mukaan 10 vuoden kuvaajia. Uutena kiinnostavana tautiryhmänä nousi esiin infektioaudit, johon olisi kiinnitettävä huomiota jatkossa tehtävissä julkaisuissa. Vaikka myös altisteet, ammatit ja toimialat ovat kiinnostavia tietoja, on selvää, että läheskään niitä kaikkia tai niiden kombinaatioita ei voida yhdessä vuosijulkaisuissa kovin tarkasti tuoda esille. Siispä jatkossa olisi kannattavaa tehdä myös kattavampia pidemmän aikavälin, kuten esimerkiksi viiden vuoden katsauksia.

Avainsanat: ammattitaudit, ammattitautiepäilyt, ammattitautijulkaisu, rekisteri, tiedon käyttäjä, tietopalvelupyyntö, TPSR

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

# Sisällysluettelo

1	JOHDANTO .....	1
2	TEOREETTINEN VIIITEKEHYS.....	3
2.1	Työperäisten sairauksien rekisteri .....	3
2.2	Tiedon kertyminen rekisteriin .....	4
2.3	Rekisteriaineisto .....	6
2.3.1	Onko rekisteri asiakirja-aineistoa? .....	6
2.3.2	Onko rekisteri tutkimusdataa? .....	9
2.4	Julkaisut ja tietopalvelupyynnöt rekisteristä.....	12
3	KIRJALLISUUSKATSAUS .....	14
3.1	Käyttäjryhmittely .....	15
3.2	Tietokäyttäytyminen ja tiedonhankinta .....	17
3.2.1	Datan tulkinta, ymmärtäminen ja kelvollisuus .....	19
3.2.2	Ammattilaisten tiedonlähteet.....	21
3.3	Rekisteritutkimus.....	22
3.3.1	Laadukas rekisteri.....	26
4	TUTKIMUSASETELMA .....	28
4.1	Tutkimuskysymykset ja tutkimusmenetelmä .....	28
4.2	Rekisterin tietosisältö .....	30
4.3	Tutkimuksen toteutus .....	32
5	TULOKSET.....	38
5.1	Tietopalvelupyyntöjen tekijät ja määrät .....	38
5.2	Tietopalvelupyyntöjen hakuehdot eli muuttujat.....	43
5.2.1	Tautiryhmä eli sairaus .....	45
5.2.2	Altiste.....	46
5.2.3	Toimiala.....	47
5.2.4	Ammatti .....	49
5.2.5	Pyyntöjen vuosijakauma.....	50
5.3	Pyyntöjen sisältö käyttäjryhmittäin .....	51
5.3.1	Opiskelijoiden pyynnöt.....	51
5.3.2	Erikoistuvien lääkäreiden pyynnöt .....	51
5.3.3	Työssäkäyvien pyynnöt .....	52
6	JOHTOPÄÄTÖKSET JA ARVIOINTI .....	53
	LÄHTEET .....	59

# 1 JOHDANTO

Tieteen termipankki (2020a) määrittelee avoimen tieteen osalta rekisterin olevan ”järjestelmä, johon julkisella sektorilla syntyneitä, viranomaistoiminnan yhteydessä kattavia tietoaaineistoja kerätään”. Avoimen tieteen pyrkimyksenä on muun muassa edistää näiden tietoaaineistojen avointa dokumentointia. Työperäisten sairauksien rekisteri eli TPSR sisältää tietoa suomalaisten ammattitaudeista ja ammattitautiepäilyistä. Rekisteri on perustettu Suomen ammattitautirekisterinä vuonna 1964. Nykymuodossaan rekisteri on laaja tietokanta, jonne päivitetään vuosittain tietoa virallisten tilastonpitäjien lähteistä. TPSR on tutkimusrekisteri, joka lain mukaan saa tietoa virallisilta rekisterinpitäjiltä, sekä tarkentavaa tietoa aluehallintoviranomaisilta. Rekisteriä hallinnoi rekisteritiimi, joka toimii Työterveyslaitoksella. Työterveyslaitos kuuluu Sosiaali- ja terveysministeriön alaisuuteen.

Työperäisten sairauksien rekisteristä julkaistaan rekisteröintivuosittain Suomen virallinen tilasto ammattitaudeista ja ammattitautiepäilyistä eli niin kutsuttu vuosijulkaisu. Rekisteristä luovutetaan tietoa myös esimerkiksi tilastojen loppukäyttäjille tietopalvelupyyntöjen perusteella, sekä alan tutkimuksiin. Rekisterin tiedot vahvistetuista ammattitaudeista luovutetaan vuosittain Eurostatille, joka on Euroopan komission alainen tilastotiedon tuottajayksikkö. Näiden lisäksi rekisterin tietoja julkaistaan verkossa Työelämä-tieto-sivustolla, joka on Työterveyslaitoksen kehittämä verkkopalvelu.

TPSR:n julkaisujen ja etenkin niiden kehittämisen osalta rekisteritiedon loppukäyttäjät ovat jääneet taka-alalle. Vaikka rekisteristä on sen olemassaolon alusta saakka julkaistu vuosittaisia katsauksia, ei koskaan ole pohdittu tarkemmin, millaisia loppukäyttäjiä tiedolla on, ja mihin tarkoituksiin rekisterin tietoa käytetään. Toimin Työperäisten sairauksien rekisterin datan laadusta ja teknisistä asioista vastaavana erityisasiantuntijana, joten

tämän pro gradun aihe ja tutkimusongelma heräsi päivittäisestä työstä sekä rekisterin julkaisujen kehittämistarpeista. Tavoitteena on, että tutkimuksen myötä rekisterin käyttöä ymmärretään paremmin ja jatkossa voidaan sen perusteella suunnitella rekisteristä tuotettavia julkaisuja paremmin.

Tutkielma rakentuu johdannon lisäksi teoriaosuudesta, jossa esitellään Työperäisten sairauksien rekisteri, selvitetään tiedon kertyminen rekisteriin ja rekisteristä tehtävät julkaisut, sekä pohditaan rekisteriaineiston olomuotoa. Kirjallisuuskatsauksessa tarkastellaan tutkimuksia, jotka liittyvät tiedon käyttäjiin, tietokäyttäytymiseen ja tiedonhankintaan. Varsinaisia rekistereiden käyttöön liittyviä rekisteritutkimuksia löytyy kirjallisuudesta hyvin niukasti, ja tarkastelussa ovat pääsääntöisesti arkistojen käyttö ja käyttäjät. Ne rekisteritutkimukset on huomioitu, jotka ovat tämän pro gradun kannalta olennaisia. Neljännessä luvussa esitellään tutkimusasetelma sekä tutkimuksen toteutus. Viides luku kokoaa tutkimuksesta saadut tulokset yhteen, ja kuudes eli viimeinen luku keskittyy tulosten arviointiin ja niistä saatuihin johtopäätöksiin.

## 2 TEOREETTINEN VIITEKEHYS

Tässä luvussa esitellään Työperäisten sairauksien rekisteri ja prosessi, miten tieto kertyy rekisteriin. Rekisteriaineiston olemusta pohditaan syvemmin, ja otetaan kantaa, onko kyseessä asiakirja-aineistoa vai tutkimusdataa. Omana osanaan tarkastellaan rekisteristä tehtäviä julkaisuja, tietopyyntöjä ja EU:lle tehtäviä tilastoja.

Dataa voidaan pitää informaation raaka-aineena. Informaatiotutkimuksen mukaan data muuttuu informaatioksi silloin, kun sen järjestys muuttuu mielekkääksi. (Haasio, Harvainen & Savolainen. 2019, 16, 20-21). Tässä pro gradu -työssä puhutaan tietokannassa olevasta *raakadatasta*, joka on vielä käsittelemätöntä informaatiota, ennen kuin siitä työtetään tiedon tarvisijalle mielekästä tilastotietoa. Tilastollinen tieto puolestaan kertoo tarkemmin jonkin asian määrällisestä luonteesta (Niiniluoto 1996, tässä Haasio ym. 2019, 19).

### 2.1 Työperäisten sairauksien rekisteri

Työterveyslaitos ylläpitää muutamia lakisääteisiä henkilörekistereitä. Kun rekisterin tiedoissa on mukana luonnollisen henkilön tunnistetietoja, on kyseessä henkilörekisteri. Henkilörekistereihin ja niiden käyttöön sovelletaan henkilötietolakea ja EU:n yleistä tietosuoja-asetusta. Asetuksessa säädetään muun muassa henkilötietojen käsittelystä ja informointivelvoitteesta. Tietosuojalainsäädäntö määrittää tietosuojaperiaatteet, joiden mukaan henkilötietojen käsittelyä on noudatettava. (Työterveyslaitos 2020, Tietosuojavaltuutetun toimisto 2020b.) Euroopan tietosuojaneuvosto on vastuussa EU:n yleisen tietosuoja-asetuksen soveltamisesta jäsenmaissaan (Tietosuojavaltuutetun toimisto 2020a).

Työperäisten sairauksien rekisteri, johon viitataan tässä työssä myös lyhenteellä *TPSR*, on Työterveyslaitoksen ylläpitämä henkilö- ja potilastietoa sisältävä tutkimusrekisteri. Laki Työterveyslaitoksen toiminnasta ja rahoituksesta (159/1978 2 a § 2 mom.) määrittelee rekisterin pidon työstä johtuvien sairauksien tutkimuksen, selvittelyn, ehkäisyyn, diagnosoinnin ja hoidon kehittämisen perusteeksi. Virallisina rekisterinpitäjinä toimivat Tapaturmavakuutuskeskus (TVK) ja Maatalousyrittäjien eläkelaitos (Mela), joiden tietojen luovutuksen TPSR:ään määrittävät Työtapaturma- ja ammattitautilaki (459/2015) sekä Maatalousyrittäjän työtapaturma- ja ammattitautilaki (873/2015). (Koskela ym. 2020, 87.)

TPSR on perustettu 57 vuotta sitten, vuonna 1964. Rekisteriin oli kirjattu vuonna 2021 noin 260000 tapausta. Rekisterin varhaisimpina vuosina tapaukset kirjattiin käsin ja rekisteriä pidettiin analogisesti kynän ja paperin avulla. Tietotekniikan kehittyessä rekisteri sähköistyi, ja on vuosien saatossa ollut monen eri palveluntarjoajan tietokannassa ja serverillä.

## **2.2 Tiedon kertyminen rekisteriin**

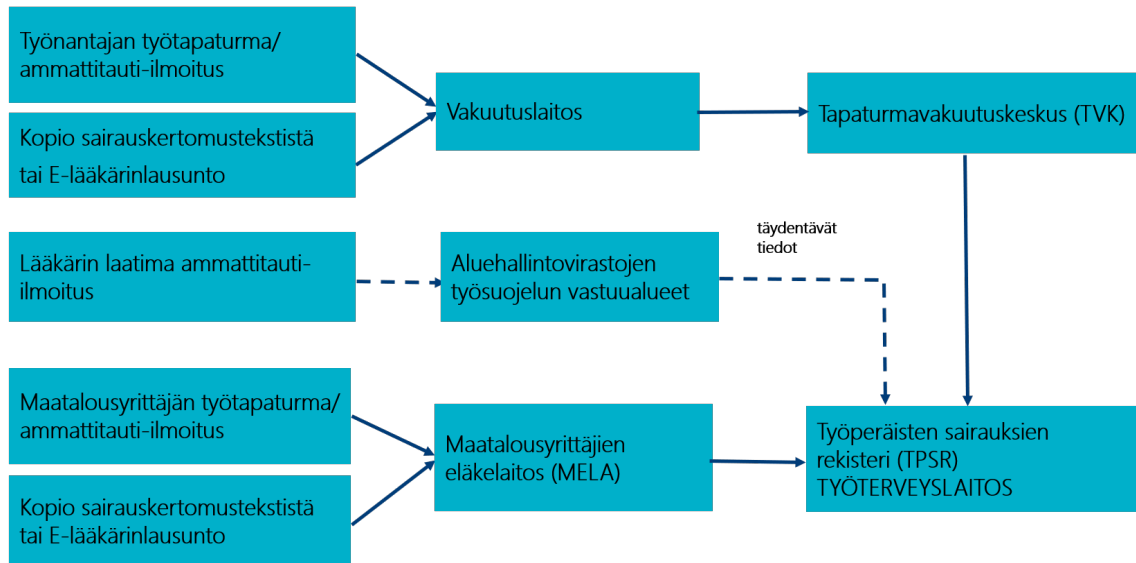
Uutta tietoa ammattitaudeista ja ammattitautiepäilyistä kertyy Työperäisten sairauksien rekisteriin vuosittain noin neljä tuhatta tapausta. Varsinkin vahvistettujen ammattitautien määrä on laskussa, mutta tätä lukumäärää voidaan pitää keskiarvona (Koskela ym. 2020, 3). Uusi data saapuu kerran vuodessa sähköisesti Tapaturmavakuutuskeskuksesta sekä Maatalousyrittäjien eläkelaitoksesta. TVK:lle tiedot ovat alun perin tulleet Suomen kaikista muista vakuutuslaitoksista paitsi Melasta, ja TVK on ensin käsitellyt tietoja omassa järjestelmässään. Kun TVK:n ja Melan lähettämät vuosittaiset datatiedostot yhdistetään TPSR:ssä, eli käsitellään yhteneväisiksi sisällön osalta ja laadullisesti mahdollisimman

hyviksi, saadaan kokoon vuosittainen aineisto ammattitaudeista ja niiden epäilyistä. Kuvassa 1 on selvitetty, miten tieto kertyy rekisteriin.

Työperäisten sairauksien rekisterillä on oikeus saada tietoja myös Aluehallintovirastojen työsuojelun vastuualueilta. Tämän määrittää laki työsuojelun valvonnasta ja työpaikan työsuojeluyhteistoiminnasta (44/2006). Nämä tiedot ovat lääkärin tekemiä ilmoituksia ammattitaudeista ja ammattitautiepäilyistä sekä työstä johtuneista muista sairaalloisista tiloista, ja ovat niin sanottuja täydentäviä tietoja. Täydentävät tiedot eivät yksinään ole rekisteriin kerättävää varsinaista dataa, vaan ne täsmentävät virallisilta rekisterinpitäjiltä tulleita tietoja esimerkiksi diagnoosien ja altisteiden osalta. (Koskela ym. 2020, 87.)

Rekisterin tietoja käsittelevät rekisterin tekninen vastuuasiantuntija ja tietojen lääketieteellisestä oikeasta laadusta huolehtiva vastuulääkäri. Rekisterin tietosisältö on monimutkainen kokonaisuus. Tätä kokonaisuutta selvitetään tarkemmin luvussa 4.2. Jotta rekisteriin tuleva data saadaan käsiteltyä mahdollisimman laadukkaaksi ja yhteneväiseksi, täytyy vastuuhenkilöillä olla oman alansa osaamisen lisäksi laaja tuntemus rekisterin sisälöstä. Datan käsittelyn jälkeen siitä tehdään kulloisenkin kalenterivuoden vahvistettujen ammattitautien ja ammattitautiepäilyjen vuosijulkaisu. Vuosijulkaisun tekemiseen kuuluu useita tilastoajoja, joiden pohjalta rakennetaan julkaisun teksti- sekä kuvaaja- ja taulukko-osuus.





Kuva 1. Tiedon kertyminen Työperäisten sairauksien rekisteriin (Koskela ym. 2020, 88).

## 2.3 Rekisteriaineisto

Työperäisten sairauksien rekisteri sisältää paljon dataa ja monia eri muuttujia. Rekisterin voi kuvailla olevan tutkimusdataa sisältävä tietovarasto, joka pitää sisällään asiakirja-aineistoa. Tarkkaa tai täydellistä jakoa aineiston olomuodosta ei ole tarpeen tehdä. Tässä luvun tarkoituksena on pohtia tarkemmin rekisterin sisältöön liittyviä näkökulmia.

### 2.3.1 Onko rekisteri asiakirja-aineistoa?

Asiakirjoilla voi olla niiden elinkaaren aikana useita eri konteksteja ja käyttäjäryhmiä, joten kyseessä on moniselitteinen käsite (Henttonen 2015, 15). Se, missä ympäristössä asiakirja on, voi muuttaa näkökulmaa paljonkin, ja esimerkiksi lainsäädännöllisesti vain hyvin pieni osa asiakirjahallinnan aineistosta saattaa olla kontekstista riippuen varsinaisia asiakirjoja. (Henttonen 2015, 16-17.) Asiakirjan tärkeimpänä tuntomerkkinä pidetään sen yhteyttä toimintaan niin organisaatioiden kuin yksilön tasolla. Vaikka asiakirja ei olisi

enää ajankohtainen, voidaan sen avulla silti suunnitella uutta ja ratkoa ongelmia. Tieteellinen tutkimus hyötyy asiakirjoista, jopa useiden vuosien päästä niiden laatimisesta, myös TPSR:n aineistosta. Työperäisten sairauksien rekisterin aineisto on kertynyt lakisääteisen tehtävän myötä ja viranomaisen voi käyttää sitä tarpeen mukaan myös päätöksenteossa. (Lybeck ym. 2006, 13-14.)

Ympäristöjen sähköistymisen seurauksena on tullut pulmia varsinaisten asiakirjojen tunnistamisessa ja erottamisessa muusta tietoaineistosta ja datasta. Enää ei voida esimerkiksi ulkoisten seikkojen perusteella eritellä asiakirjoja ja muuta aineistoa, kun kaikki dokumentit saattavat näyttää digitaalisesti samanlaisilta. (Henttonen 2015, 17-18.) Varsinainen rekisteri sekä suurin osa muusta TPSR:n aineistosta on sähköisenä, mutta osa täydentävistä tiedoista on edelleen paperimuotoista aineistoa, jota ei ole sähköistetty. Rekisterin koko sisältö ei siis ole yksi tiettyä muotoa oleva aineisto, vaan siihen kuuluu erilaista materiaalia, joka on tullut eri lähteistä. Silti molemmissa tapauksissa koko aineiston ja yksittäisten asiakirjojen merkitys on tärkeä.

Aikaisempina vuosina TPSR:ään lähetettiin täydentäviä tietoja eli ammattitauti-ilmoituksia paperiaineistona, jotka arkistoitiiin lukollisiin kaappeihin. Nykyisin tämä aineisto on sähköistetty jo lähettävän tahon toimesta. Kyseinen sähköinen aineisto tallennetaan potilastietojärjestelmään, jotta henkilötietoja sisältävien asiakirjojen tietoturva ja tietosuojasäilyvät. Asiakirjan tärkeinä piirteinä pidetään luotettavuutta, autenttisuutta, eheyttä ja käytettävyyttä (Henttonen 2015, 18-21). TPSR:ään tulevan aineiston osalta nämä asiakirjan piirteet toteutuvat. Silti jos pohditaan, voidaanko kaikkea rekisterin sisältämää dataa ja tietoa pitää nimenomaan asiakirjallisena aineistona, niin mielestäni vastaus on kuitenkin ei.

Julkiset asiakirjat voidaan jakaa kahteen eri arvoluokkaan. Primäärin arvon saavat ne asiakirjat, jotka ovat luomisorganisaatiossaan, ja sekundäärin arvoisia ovat eri organisaatioissa ja muualla kuin virallisessa julkisessa käytössä olevat asiakirjat (Schellenberg 1956, 28). Muitakin vastaavia jakoja ovat miettineet muun muassa McKemmish (1993), joka arvotti asiakirjoja esimerkiksi niiden luojien ja käytön mukaan, sekä Smedberg (1995), jonka mukaan asiakirjat oli analysoitava niiden käytön perusteella, eli käyttö määritteli niiden arvon (tässä Sundqvist 2007, 628-629). TPSR:n data on muodostunut eri organisaatioissa olevien primäärilähteiden datasta, eli näiden teorioiden mukaan rekisterin aineisto ei itsessään ole primääriä.

Asiakirjojen suoraa ja epäsuoraa käyttöä on tutkinut muun muassa Pugh (1992). Suora käyttö on hänen mukaansa suoraviivaista todeta. Suoraa käyttöä on esimerkiksi se, kun joku hankkii tietoa asiakirjasta henkilökohtaisesti arkistosta tai kysyy puhelimesta tai sähköpostilla. Epäsuoran käytön määrittelyminen on hankalampaa, koska käyttäjä voi saada tietoonsa asiakirjoissa olevia asioita epäsuorasti, mutta hän ei välttämättä ole niitä hankkinut henkilökohtaisesti. (Pugh 1992, 12-13.) Miten asiakirjojen suoran ja epäsuoran käytön määritelmät sitten toimivat, kun kyseessä on TPSR? Suoraksi käytöksi voidaan mieltää käyttö silloin, kun loppukäyttäjä pyytää tilastotietoja suoraan TPSR:n vastuukäyttäjien kautta. Epäsuoraa käyttöä voidaan ajatella tapahtuvan aina, kun loppukäyttäjä löytää tietoa esimerkiksi rekisterin vuosijulkaisuista, silloin kun hän löytää ne internetistä tai fyysisenä paperijulkaisuna.

TPSR:n aineisto ei käy läpi asiakirjoihin liittyvää elinkaarimallia. Rekisterissä oleva aineisto tai asiakirja on käytössä satunnaisesti aina silloin, kun kyseistä tapausta käsitellään. Tiivistettynä asiakirjan elinkaari alkaa elinkaarimallin mukaan sen laatimisesta. Tämän jälkeen asiakirja ryhmitellään tai luokitellaan, sitä käytetään aktiivisesti ja säilytetään

siihen saakka, kunnes se ei enää ole ajankohtainen (Henttonen 2015, 81-82, Tough 2006, 3). Jossain vaiheessa otetaan kantaa asiakirjan hävittämisestä, tai päätetään, että se säilytetään pysyvästi ja arkistoidaan (Tough 2006, 3). Kaikki se tieto, mikä tulee rekisteriin, myös säilytetään siellä, pois lukien duplikaattien eli niin sanottujen virheellisten tuplata-pausten poistaminen ja datan laadullinen korjaaminen.

Edellisten määritelmien ja pohdinnan mukaan rekisterin aineisto täyttäisi joiltakin osin asiakirjallisen aineiston määritelmän. Asiakirjan käsitettä voidaan kuitenkin katsoa niin monesta eri näkökulmasta, että rekisteriä määriteltäessä on mahdotonta sanoa, kumpaan suuntaan vaaka kallistuu enemmän. Mielestäni asiakirjan piirteiden tiedostaminen ja tunnistaminen, ja mahdollisesti niiden poissulkeminen ovat kuitenkin tärkein päämäärä.

### **2.3.2 Onko rekisteri tutkimusdataa?**

Tutkimusdata mielletään usein tutkijan itse tuottamaksi tai käyttämäksi resurssiksi, joka voi olla eri muodoissa olevaa aineistoa (Tieteen termipankki 2020b). Työperäisten sairauksien rekisteri on viralliselta määritelmältään *tutkimusrekisteri*. Tämä tarkoittaa sitä, että se saa tietonsa muualta, virallisilta rekisterinpitäjiltä, ja kokoaa nämä tiedot yhteen samalla kun ottaa kantaa laadullisiin asioihin, kuten kappaleessa 2.2 on selvitetty. Rekisteristä voidaan antaa tietoja esimerkiksi työterveysalan tutkimuksiin, eli tieto, joka rekisterissä on, voidaan määritellä tietynlaiseksi tutkimusdataksi. (Työterveyslaitos 2020.)

Hyvän tieteellisen käytännön ohjeistus ottaa kantaa esimerkiksi siihen, että tutkimukseen tarvittava tutkimuslupa on olemassa, ja että tutkimus suunnitellaan, toteutetaan ja raportoidaan, ja varsinkin tutkimuksessa syntyneet tietoaineistot tallennetaan vaatimusten määrittelemällä tavalla (Tutkimuseettinen neuvottelukunta 2019, Vastuullinen tiede 2020). Jos tehdään tutkimusta, jonka aineistona on henkilötietoja sisältävää materiaalia,

on sen saamiseksi ja keräämiseksi tarvittu tutkimuslupa, jotta aineisto on ylipäättään saatu (Briney 2015, 94-97). TPSR:n sisältämää dataa säilytetään lakien määrittelemällä tavalla, ja jotta sieltä voidaan luovuttaa salassa pidettävää tietoa, täytyy kyseiseen tutkimukseen olla tutkimuslupa haettuna ja myönnettynä Työterveyslaitokselta. Nykyään Suomessa tutkimusluvut eri aineistojen yhdistämiseen sosiaali- ja terveystietoja sisältäville tutkimuksille myöntää Findata, joka toimii THL:n eli Terveyden ja hyvinvoinnin laitoksen yhteydessä (Findata 2020).

Kestävään kehitykseen liittyvät asiat mielletään yleensä esimerkiksi energiaa tuottaviksi resursseiksi, kuten vesi ja metsät. Kestävät toimintatavat ovat tärkeitä siksi, että pitkällä ajanjaksolla tapahtuisi esimerkiksi talouskasvua ja että vaurastuminen olisi mahdollista. Näiden lisäksi on tärkeää ajatella myös kirjastoja ja arkistoja kestävien resurssien vaalijoina, jotta nekin säilyvät pitkään. Myös tutkimusdatat, joita täytyy säilyttää pitkiä aikoja, ja joiden arvo säilyy siksi, että niistä saadaan jatkuvasti tärkeää dataa, voidaan ajatella olevan vastaavanlaisia kestävästi vaalittavia ja säilytettäviä arvokkaita asioita. (Pryor 2012, 67-68.)

Kestävän kehityksen periaatteella hoidetaan myös Työperäisten sairauksien rekisteriä. Rekisterin dataan tehdään laadullisia korjauksia ja muun muassa poistetaan duplikaatteja, jotta sen arvo säilyisi pitkään, ja sieltä saatava tieto olisi aina luotettavaa. Tämä ajatus puoltaa TPSR:n olevan tutkimusdataa. Pryor (2012, 81) ehdottaakin, että kun on kyseessä tutkimusdata, ei puhuttaisikaan kestävästä kehityksen tutkimusdatasta, vaan kestävästä kehityksen mukaan tehtävistä datan parannustoimista.

Nykyaikainen tieteen infrastruktuuri, jota kutsutaan myös esimerkiksi nimillä kyberinfrastruktuuri ja e-tiede, sisältää tietotekniisiin mahdollisuuksiin liittyviä ajatuksia muun muassa uusista tavoista organisoida tieteellistä tietämystä ja sivistystä (Jackson ym. 2007).

Tieteen infrastruktuurin muotoutumiseen on vaikuttanut esimerkiksi *open access* eli (tiedon) avoin saatavuus. Open access on viime vuosien aikana puolttanut myös akateemisten aineistojen avointa julkaisemista. (Hughes 2012, 139.) Tutkimusaineistojen jakaminen ja virtuaalisen yhteistyön tekeminen ovat siis mahdollistuneet e-tiedekaudella (Karasti ym. 2011), mutta samalla on tunnistettu lakien aiheuttamia esteitä sekä datan, että tutkimuksen tulosten julkaisemiselle (Talja ym. 2010).

Aiempien tutkimusten mukaan vastaavanlaisia rekistereitä ja data-arkistoja, kuin Työperäisten sairauksien rekisteri, on pidetty osana tieteen infrastruktuuria (mm. Hughes 2012, 141-146). TPSR sisältää salassa pidettäviä tietoja, joten sen sisältämää dataa ei voida suoraan luovuttaa open access -tyyliseen käyttöön. Jos dataa ollaan luovuttamassa, datasta poistetaan kaikki henkilötunnisteet, ja tapaukset anonymisoidaan tai pseudonymisoidaan, eli suojataan rekisterissä olevien henkilöiden yksityisyyttä tietosuojaperiaatteiden mukaisesti. Tällaisten toimenpiteiden jälkeen dataa on mahdollista julkaista datan omistajan eli Työterveyslaitoksen päätöksellä.

Jokapäiväisissä toiminnoissa data tarkoittaa samaa asiaa kuin kyberinfrastruktuuri, mutta data voi olla paljon muutakin. Se voi olla näyte, kokoelma tai malli, se voi olla visuaalista, numeerista, tai vaikkapa osa jotain julkaisua. Kun pohditaan datan saamista, jakamista ja säilyttämistä, puhumattakaan mahdollisuuksista käyttää sitä keksintöihin, ohjelmiin ja oppimiseen, ymmärretään, että tieteen pitää myös tuottaa. Siispä data itsessään voi olla myös tuote. Nämä esimerkit osoittavat sen, että myöskään kyberinfrastruktuuri ei voi koskaan olla yksittäinen tai yhtenäinen asia. (Jackson ym. 2007.) Työperäisten sairauksien rekisteri ei varsinaisesti ole tuottava. Se voi kuitenkin välillisesti olla osa tuottavaa prosessia, jos sieltä saatava tieto auttaa myöhemmin esimerkiksi jonkin ammattitaudin tai

altisteen kattavampaa tunnistamista, sen myötä parempia suojauskeinoja työpaikoilla, ja siten yksilöiden parempaa vointia työssään.

Tutkimusinfrastruktuuripolitiikassa pitäisi tulevaisuudessa näkyä paremmin se, miten dataa voitaisiin jakaa ja esittää laajemmalla mittakaavalla. Tietokantojen yhdistäminen, niiden yhteiskäyttöisyys, datan esittäminen ja analysointi, ja myös aivan uudenlaiset tavat tuottaa tietoa olisivat tärkeitä aiheita pohtia ja tutkia enemmän. (Talja ym. 2010.) Nykyisin on alettu enenevässä määrin luoda niin sanottuja tietopankkeja, joihin yhdistetään dataa monesta eri lähteestä tai aiheesta. Yksi esimerkki on Työelämä-tieto-sivusto, jonka Työterveyslaitos on kehittänyt. Se julkaistiin vuonna 2019. Sivustolle on koottu tietoa erilaisista lähteistä, tutkimuksista ja tietokannoista. Sivustolla käyttäjä voi interaktiivisesti etsiä haluamaansa tietoa. Työperäisten sairauksien rekisteristä otettu anonymisoitu data on osa tätä sivustoa. (Työelämä-tieto 2020.) Tällaisen laajan sivuston eli palvelun kehitystyö on jatkuvaa. Jos jokin tärkeä osa, kuten esimerkiksi tekninen alusta tai sosiaalinen näkökulma unohdetaan, koko infrastruktuuri saattaa epäonnistua. (Edwards ym. 2007, 5-6.)

## **2.4 Julkaisut ja tietopalvelupyynnöt rekisteristä**

Työperäisten sairauksien rekisterin aineistosta tehdään vuosittain vuosijulkaisu, jossa esitetään ja analysoidaan tietyn vuoden vahvistetut ammattitaudit ja ammattitautiepäilyt. Nykyisin vuosijulkaisun muotoja ovat pdf-julkaisu ja graafinen eli tilastollisilla kuvaajilla tehty osio, joka tehdään ppt-tiedostona. Jotta vuosijulkaisu ei paisuisi liian suureksi ja vaikeasti hahmotettavaksi kokonaisuudeksi, on siinä pyritty esittämään tiedot joko ”top10”-tyylisinä katsauksina tai yleisimpien sairauksien mukaan esitettynä.

Joskus vuosijulkaisun tiedot eivät ole riittävän informatiivisia. Esimerkiksi tutkija voi tarvita omassa tutkimuksessaan tarkkoja tilastotietoja tutkimastaan asiasta. Vuosijulkaisussa ei voida aineiston laajuuden vuoksi analysoida kovinkaan tarkasti yksittäisiä ammattiluokkia, toimialaluokkia, sairauksia tai altisteita. Erilaisia kombinaatioita on paljon, eikä niiden esittelemisen ole tarkoituksenmukaista yleisessä vuosijulkaisussa. Näissä tapauksissa rekisteristä tehdään tietopalvelupyyntöjen perusteella myös tarkempia tietokantahakuja ja tilastoja. Tietopalvelupyyntöjen esittäjät ovat esimerkiksi eri tieteenaloilta, erilaisista yrityksistä ja yhteisöistä, mediasta sekä opinnäytetyön tekijöistä koostuvia rekisteritiedon loppukäyttäjiä.

Tilastotieto, jota TPSR:stä luovutetaan, ei sisällä henkilö- tai muita tunnistetietoja, joiden perusteella henkilön voisi tunnistaa. Rekisteristä saatava raakadata pseudonymisoidaan, jonka jälkeen siitä työstetään tilastoja, joiden sisältämä tieto on lopulta anonymisoitua. On kuitenkin olemassa tutkimushankkeita, joissa on tarve yhdistää monessa eri rekisterissä sijaitsevaa dataa. Esimerkiksi terveystietoja sisältävien rekistereiden ollessa mukana hankkeessa, tiedot luovutetaan Findatalle ja he hoitavat tietojen yhdistelyn keskitetysti.

Työperäisten sairauksien rekisterin vuosittaiset ammattitautitiedot kuuluvat virallisiin ESS:n eli European Statistical System -järjestelmän tilastoihin. Jokainen Eurostatin ammattitautitilastoinnin pilottihankkeeseen kuuluva maa (kaikki EU-maat ja muutamia EU:n ulkopuolisia Euroopan maita) valmistelee oman ammattitautitilastointiaineistonsa Eurostatin metodologian mukaan ja lähettää sen kerran vuodessa Eurostatille, joka koostaa EU:n yhteneväiset tilastot. Aineisto valmistellaan muun muassa siten, että Suomessa ammattitautitilastoinnissa käytössä olevat luokituskoodit koodataan EU:n mukaisilla koodeilla ja muuttujilla. (EUR-Lex 2021.)



### 3 KIRJALLISUUSKATSAUS

Kirjallisuuskatsauksessa tarkastellaan käyttäjätutkimuksia, varsinkin käyttäjäryhmitteilyä, sekä tiedonhankintaa ja tiedon käyttöä tiedon tarpeen ohjaamana. Koska tiedonhankintatutkimus on yksi informaatiotutkimuksen keskeisimmistä alueista (Haasio ym. 2019, 7-8), on aihetta pyritty taustoittamaan tässä työssä melko yleisellä tasolla. Varsinaisia rekistereiden käytöstä tehtyjä käyttäjätutkimuksia ei kirjallisuudesta löydy, joten fokus on arkistojen ja asiakirjojen käyttäjissä.

Tässä luvussa pohditaan myös tiedon luotettavuutta tiedon käyttäjän näkökulmasta tehtyjen tutkimusten pohjalta, sekä ammattilaisten käyttämiä tiedonlähteitä kirjallisuuden perusteella. Tarkastelussa ovat myös ne harvat tutkimukset, jotka liittyvät varsinaisten rekistereiden tutkimuksiin.

Täysin vastaavanlaista tutkimusta kuin tämän pro gradun aihe ei kirjallisuudesta löytynyt, mutta tietopalvelupyynnöiden hakutermejä (Mäkivaara 2017) ja tutkimusaineistojen metatietoja (Pajari 2019) on graduissa analysoitu. Yksityisen arkiston asiakirjoihin kohdistuneissa tietopalvelupyynnöissä olleita hakutermejä tutkittiin tapaustutkimuksen keinoin, ja tarkoituksena oli olemassa olevan metatietomallin hyödyntämisen parantaminen tai uuden mallin kehittäminen (Mäkivaara 2017). Tutkimusaineistojen metatietojen kirjaaminen puolestaan on edellytys tutkimusaineistojen avoimuudelle, eikä vastaavaa metatietojen laadun kartoitusta ollut ennen Suomessa tehty data- ja metatietoarkistoissa. Tutkimus tehtiin tilastollisin menetelmin ja tuloksia tarkasteltiin muun muassa tunnuslukujen avulla. (Pajari 2019.)

### 3.1 Käyttäjärhmittely

Informaatiotutkimuksessa sekä tieto- ja asiakirjahallinnan alalla on tehty paljon tutkimuksia käyttäjistä. Käyttäjiä on tutkittu monista eri lähtökohdista esimerkiksi tiedon etsijöinä, käyttäjinä ja jakajina. Nykyään ymmärretään, että arkistojen käyttäjien ymmärtäminen ja analysointi voivat auttaa palvelujen parantamisessa monille erilaisille käyttäjäryhmille (Sundqvist 2009, 12). Belkin (1984) kuvaa ihmistä, joka etsii tietoa, tiedon käyttäjäksi, joka on samalla tiedon tarvitsija, ja joka hyödyntää saamaansa informaatiota (Haasio ym. 2019, 32). Digitalisaatio ja sähköiset järjestelmät ovat laajentaneet erilaisten arkistojen käyttäjäkuntaa, ja sen myötä käyttäjämäärän tutkimukset ovat lisääntyneet (Hughes 2012, 3).

Asiakirjojen käytön motivaatiosta ja tarpeesta on tehty useita tutkimuksia, mutta ei niinkään asiakirjojen varsinaisesta käytöstä (Sundqvist 2007, 627). Myös sitä on tutkittu, millaisiin tarpeisiin asiakirjoja on tarvittu. Menneisyudessa asiakirjojen käyttäjät olivat usein historioitsijoita, ja asiakirjojen tarve koski heidän tutkimuksiaan. Myös politiikkaa, kirjallisuutta ja taiteen tutkimusta varten on asiakirjoja etsitty ja käytetty. (Smedberg 1980; tässä Sundqvist 2009, 8.) Tutkimuksia on tehty myös opettajien asiakirjojen käytöstä (mm. Gilliland-Swetland 1998), sekä kehittyvien maiden tarpeista käyttää asiakirjoja, mutta enemmän kuitenkin siellä olevista asiakirjahallinnon pulmista (mm. Kemoni ym. 2003). (tässä Sundqvist 2009, 10-11.)

Pugh (1992) jaottelee arkistojen ja asiakirjojen käyttäjät kahteen eri pääryhmään: ammatilliset käyttäjät ja harrastelijakäyttäjät. Pääryhmien sisällä hän määrittelee tarkemmin eri alaryhmiä. Ammatillisia käyttäjiä ovat esimerkiksi organisaation henkilöstö,

lakiasiantuntijat, insinöörit, tutkijat, opiskelijat ja opettajat. Harrastelijakäyttäjiksi luetaan sukututkijat ja historiaa harrastuksenaan tutkivat henkilöt. (Pugh 1992, 14-24.)

Ford puolestaan on luokitellut erilaiset kontekstit, jotka vaikuttavat yksilöiden tietokäyttämiseen. Näitä ovat työ, opiskelu ja vapaa-aika. (Ford 2015, 100.) Fordin ja Pughin jaottelut voisi limittää toisiinsa siten, että Fordin työ ja opiskelu ovat alueita, joihin kuuluvat Pughin luokituksen ammatilliset käyttäjät. Pughin harrastelijakäyttäjät voisi luokitella puolestaan vapaa-ajan tietokäyttämisen puolelle Fordin luokituksessa. Aina jaottelua ei kuitenkaan voi tehdä kovin suoraviivaisesti, ja käyttäjäryhmittelyyn liittyy myös monia yksilöön vaikuttavia sisäisiä ja ulkoisia tekijöitä, jotka täytyy ottaa huomioon (Ford 2015, 102).

Ammattiryhmät ovat ensimmäinen rakenteellisesti tutkittu, ja myös ylipäättään eniten tutkittu sosiaalinen käyttäjäryhmä tiedonhankintakäytäntöjen tutkimuksessa. Tutkimukset on aloitettu tutkijoista, insinööreistä, sosiaalitieteilijöistä ja humanisteista. Sen jälkeen on tutkittu muita ammatteja ja ihmisten arkisia rooleja. Tutkittuja arkirooleja ovat olleet esimerkiksi kuluttaja, potilas ja vanhempi. Erilaisten ryhmien tietokäyttämisen on tehty monia vaihtelevia tutkimuksia. Tutkimuksissa on havaittu tutkijoiden yhteneväinen huoli lähteiden ja kanavien laadusta, kun ihmiset etsivät informaatiota. Ihmiset luottavat yhä toisiin ihmisiin esimerkiksi tiedon lähteenä, vaikka kaikenlaista tietoa on kenen tahansa helppo jakaa nykyään verkossa. (Case & Given 2016, 278-279, 320, 345-346.)

Työperäisten sairauksien rekisteri sisältää tietoa ammattitaudeista ja niiden epäilyistä, joten rekisteristä pyydetään tietoa silloin, kun halutaan tilastollista tietoa näihin liittyvistä asioista (Koskela ym. 2020, 87). Tämä on melko tarkka ja rajattu aihe. Kuitenkin tiedon tarve ammattitautitilastojen osalta tarkentuu jokaisen tietopalvelupyynnön kohdalla enemmänkin. Usein arkistojen kohdalla on tutkittu, miten jokin tietty ammatti- tai muu

ryhmä käyttää tai etsii asiakirjoja ja informaatiota, joten tässä työssä tutkimus kääntyy toisin päin. Nyt haluttiin selvittää, mitkä ammatti- tai muut ryhmät informaatiota pyytävät, kuinka paljon, ja mitä informaatio sisältää.

### **3.2 Tietokäyttäytyminen ja tiedonhankinta**

Yleisellä tasolla tiedonhankinta voidaan ajatella toiminnaksi, joka alkaa jostakin tiedontarpeesta, ja sen olennaisena osana on relevanttien tiedonlähteiden ja kanavien löytäminen ja hyödyntäminen. Tiedonhankinnan prosessissa on aina olemassa jokin päämäärä, kuten esimerkiksi opiskelun tai ongelmanratkaisun edistäminen. (Haasio ym. 2019, 35.)

*Tiedontarve* (engl. information need), saa yksilön hankkimaan tietoa jostain asiasta, josta ei vielä tiedä tarpeeksi (Ford 2015, 18). Yksilö on tietoinen siitä, että tarvitsee jotakin informaatiota, jolla voi ratkaista jonkin ongelman (Haasio ym. 2019, 39). Tarve voi liittyä myös haluun oppia lisää jostain aiheesta, ja yksilö ryhtyy kyseisen tiedontarpeen myötä etsimään tietoa asiasta (Ford 2015, 29). Tiedontarve voi olla aluksi epämääräistä, ja sitten selkiytyä prosessin kuluessa tietoisemmaksi ja jäsentyneemmäksi tarpeeksi (Haasio ym. 2019, 65-66). Tiedontarve ja tiedon etsiminen ovat osa laajempaa kokonaisuutta, jota kutsutaan *tietokäyttäytymiseksi* (engl. information behavior). Joskus siihen voi kuulua myös halu välttää tietynlaista informaatiota. (Ford 2015, 49.)

Yksilön tietokäyttäytymisen ymmärtäminen on tärkeää monesta syystä. Nykyään informaatioympäristömme on enenevässä määrin hajanainen (Johnson 2014, 715). Jotta voitaisiin parantaa yksilön tai yhteisön pääsyä oikean tiedon pariin, täytyy pyrkiä ymmärtämään paremmin tiedon hakijan ja tiedon tarvitsijan tietokäyttäytymistä. Kun tätä ymmärretään, voidaan vaikuttaa moneen muuhunkin asiaan, kuten esimerkiksi siihen, miten tiedon tuottaja pystyy esittämään tiedon paremmin ja tehokkaammin juuri kulloinkin

kyseessä olevalle tiedon tarvitsijalle. (Ford 2015, 8.) Yksinkertaistettuna voidaan todeta, että esimerkiksi ohjelmistojen kehitystyö paranee, kun kehittäjä pystyy suunnittelemaan älykkäämpiä työkaluja sen perusteella, mitä tiedetään käyttäjien tietokäyttäytymisestä. Älykkäämmän ohjelman käyttäminen puolestaan voi auttaa ihmistä saamaan tarvitsemansa tiedon helpommin, ja näin välillisesti auttaa tiedon tarvitsijaa tulemaan älykkäämmäksi. (Ford 2015, 217.)

Yhtenä tärkeänä osana Työperäisten sairauksien rekisterin vastuuhenkilön työtä on kommunikointi tietopalvelupyynnön tekijän kanssa, jotta isosta tietokannasta ja suuresta tietomäärästä pystytään etsimään ja löytämään juuri se tietty haluttu tulos. Joskus eivät tiedon pyytäjien tiedontarpeet ole heille itselleenkaan aluksi selkeitä, vaikka tiedon käyttötarkoitus olisikin, joten rekisterin vastuuhenkilön on osattava kysyä järkeviä tarkentavia kysymyksiä, jotta päästään haluttuun lopputulokseen. Tiedon tarve saattaa muotoutua prosessin ja kommunikoinnin myötä, ja kysymysten muotoilu ja termien käyttö saattaa johtaa tiedon tarpeen uudelleenmäärittelyyn (Savolainen 2012).

Varsinaisten asiakirjojen etsiminen saattaa olla työläs prosessi. Sundqvist (2009) selvitti, millä tavalla henkilö tekee hakuja, käyttääkö hän hakutoimintoja vai esimerkiksi oikeaa ihmistä apuna. Tämän kirjallisuuteen pohjautuvan selvityksen tulokset olivat monitahoiset. Määrittäviä tekijöitä siihen, miten hakuja tehtiin, olivat muun muassa henkilön ammatti, henkilön ymmärrys järjestelmien teknisistä termeistä, aiempi tietämys kyseisestä aihealueesta, sekä se, tunsiko henkilö ennestään arkistoammattilaisen, jolta pystyi henkilökohtaisesti pyytämään apua. (Sundqvist 2009, 14-20.) Sundqvist toteaa, että koska kirjallisuudesta löytyvät tutkimukset asiakirjoja etsivien tietokäyttäytymisestä ovat yleensä toisiinsa verrattuna tehty aivan eri näkökulmista ja toteutettu erilaisilla metodeilla, on niiden pohjalta vaikeaa tehdä vertailuja tai yleistyksiä (Sundqvist 2009, 23, 222).

### 3.2.1 Datan tulkinta, ymmärtäminen ja kelvollisuus

Asiakirjojen käyttöä, niiden etsimis- ja löytämisprosessia, on tutkittu vain vähän. Asiakirjan yksi tärkeä ominaisuus on, että se voidaan luoda aivan eri paikassa ja kontekstissa, kuin missä sitä myöhemmin saatetaan käyttää. Myös asiakirjan luoja ja käyttäjä ovat usein eri henkilöitä. Luonteeltaan asiakirjat ovat monimutkaisia, ja jo se, että oikean asiakirjan pystyy löytämään oikeasta paikasta, vaatii taitoa etsiä, tulkita ja ymmärtää kontekstia ja sisältöä oikein. (Sundqvist 2009, 1-2.) Hankitun tiedon hyödyntämistä kutsutaan suppeasti *tiedonkäytöksi* (engl. information use), joka on eri asia, kuin tiedon varsinainen soveltaminen (Haasio ym. 2019, 42).

Varsinaisen tiedon tai datan esittäminen tiedon tarvitsijalle on tärkeä asia tehdä oikein. Sellaiselle, joka ei ymmärrä datan kontekstia, data voi olla ainoastaan lista merkityksettömiä numeroita tai tekstiä. Tarvitaan joku, joka osaa tulkita dataa, ja tulkinnan perusteella pystyy muuttamaan sen ymmärrettäväksi tiedoksi. (Ford 2015, 13.) Datan tulkinna on tärkeää myös oikeanlainen kognitiivinen auktoriteetti, eli ymmärrys siitä, että tiedon tulkitsijaan voi luottaa (Wilson 1983, tässä Haasio ym. 2019, 26). Varsinaisen tiedon organisointi ja hallinnointi ovat tärkeitä vaiheita sekä tiedon pyytäjälle, että organisaatiolle, joka tiedon luovuttaa (Ford 2015, 22). Informaation jakaminen on kuitenkin aina yhteistyötä, ja sillä myötävaikutetaan jonkin tuloksen aikaansaamiseksi tai pulman ratkaisemiseksi (Ford 2015, 77, 81).

Jokainen yksilö käyttää henkilökohtaisessa oppimisprosessissaan tietyn tyyppisiä keinoja. Näitä samoja keinoja käytetään myös informaation etsimisessä, jos esimerkiksi halutaan ratkaista jokin pulma. Joku haluaa aina aloittaa siten, että luo kattavan yleiskuvan asiasta. Joku toinen puolestaan hahmottaa asiat pala palalta, ja etsiminen etenee samalla. Fordin (2015) mukaan molemmilla tyyleillä on mahdollista päästä samaan

lopputulokseen. On kuitenkin havaittu, että erilaiset tyyli saattavat johtaa siihen, että asiat ymmärretään eri tavoin, ja tästä johtuen tiedon syvälinen ymmärtäminen ei olekaan täysin samanlaista. Kun on tutkittu nuorempia sukupolvia (syntyneet 1993 jälkeen), on heidän tietokäyttäytymisessään (tiedontarpeissa, tiedon haussa) havaittu tietynlaisia puutteita. Ensinnäkin he etenevät oletusten kautta. Toiseksi he haluavat löytää tiedon mahdollisimman pian, mahdollisimman pinnallisella eli vähäisellä etsinnällä, eivätkä välitä, onko tieto luotettavaa, tai että jääkö jokin oleellinen tieto huomaamatta. (Ford 2015, 59-61.)

Tietokäyttäytymisen yksi tärkeä osatekijä on se, miten hyvin pystytään arvioimaan saadun tiedon kelvollisuus suhteessa tiettyyn tiedon tarpeeseen. Eli onko saatu tieto tarpeeksi relevanttia, luotettavaa tai ymmärrettävää. (Ford 2015, 20.) Ford kuitenkin muistuttaa informaatioammattilaisia, että asiakasta, varsinkaan nuorta, ei saa kohdella alentuvasti, vaikka huomaisi hänen kriittisen ajattelunsa puutteellisuuden suhteessa löydetyn informaation laatuun (Ford 2015, 73-74). Kaikista merkityksellisimmän tiedon tunnistaminen tietylle ihmiselle tietyllä hetkellä, on huomattu olevan monimutkainen prosessi (Case & Given 2016, 110).

On monia sisäisiä ja ulkoisia tekijöitä, jotka vaikuttavat tietokäyttäytymiseen. Esimerkiksi yksilön oma kognitiivinen tyyli vaikuttaa siihen, miten hän etsii ja ymmärtää informaatiota. Kognitiivisia tyylejä on tutkittu paljon oppimisen ja koulutuksen osalta, mutta niitä on alettu ottaa huomioon enenevässä määrin myös tietokäyttäytymisen tutkimuksessa. Ulkoisista tekijöistä esimerkiksi organisaation oma sisäinen tyyli, eli se, miten asiat hoidetaan työpaikalla, saattaa vaikuttaa paljonkin yksilön tietokäyttäytymiseen. (Ford 2015, 106, 120.)

Kulttuurisia ilmiöitä ei myöskään voida väheksyä tiedonhankinnasta puhuttaessa. Kulttuurisilla tekijöillä on suuri vaikutus tiedonhankintaan sekä työ- että opiskeluyhteisöissä.

Kulttuurisidonnaisia eroja on huomattu esimerkiksi siinä, millaisia lähteitä eri maiden ihmiset käyttävät. (Haasio ym. 2019, 75-76.)

Koska tieto itsessään ei käyttydy, on tietokäyttäytyminen käsitteenä ongelmallinen. Informaatiotutkimuksessa onkin alettu käyttää enemmän käsitettä *tietokäytännöt* (engl. information practices), sillä sen ei oleteta tutkivan käyttäytymistä syvällisesti, vaan enemmänkin sosiaalisia ja kulttuurisia vakiintuneita tapoja, joilla yksilö tunnistaa tiedonlähteitä ja kanavia. (Savolainen 2008, tässä Haasio ym. 2019, 46-47.)

### **3.2.2 Ammattilaisten tiedonlähteet**

On tehty tutkimuksia siitä, millaisia lähteitä ammattiryhmät käyttävät pääasiallisesti tiedon hankintaan. *Tiedonlähde* määritellään informaatiotutkimuksessa halutun informaation sisältämäksi informaation kantajaksi. Sellaisia ovat esimerkiksi dokumentit tai muut ihmiset. (Haasio ym. 2019, 28-29).

Borglund (2005) tutki kolmen kvalitatiivisen tutkimuksen avulla, millä tavoin poliisit käyttävät sähköisiä asiakirjoja ja arkistoja operationaalisessa toiminnassa Ruotsissa, ja kuinka nämä vaikuttavat heidän työhönsä. Tulokset osoittivat, että poliisin tietojärjestelmät, jotka sisälsivät sähköisiä asiakirjoja, olivat tärkeimmät lähteet tiedon etsintään, ja niitä käytettiin lähes aina ensimmäisinä. Vasta sitten, jos näistä ei löytynyt informaatiota, jota sillä hetkellä oli tarpeen löytää, alettiin tietoa etsiä muualta. Tutkimukseen osallistuneet poliisit olivat suurelta osin samaa mieltä siitä, että informaatio ja tieto ovat samanarvoisia asioita, koska näitä tarvitaan taktisesti ja laillisesti oikeiden päätöksien tekemiseen. Ja luonnollisesti mitä enemmän informaatiota on saatavilla esimerkiksi vakavan tilanteen ollessa päällä, sen arvokkaampi ”ase” tieto on. (Borglund 2005.)



Tiedonhankintaa tehdään nykyään yleisesti verkossa (Haasio ym. 2019, 30-31), ja myös Työperäisten sairauksien rekisteristä tietoja pyytävät henkilöt kertovat usein, että ovat yrittäneet etsiä haluamaansa tietoa ensin verkosta. Informaatiotutkimuksessa puhutaan *vähimmän vaivan periaatteesta* (Zipf 1949), joka nimensä mukaisesti tarkoittaa ihmisen etsivän tietoa sieltä, mistä olettaa sen vaivattomimmin saavan. Nykyään tämä tarkoittaa siis internetiä, eikä niinkään enää esimerkiksi analogisia painettuja lähteitä. (Haasio ym. 2019, 31.) TPSR:ään yhteyttä ottaneet tiedon tarvitsijat ovat kertoneet, että he ovat monesti törmänneet seikkaan, että juuri vähäisimmällä vaivalla löytynyt tieto on vaikuttanut vanhentuneelta. Tämän vuoksi he ovat alkaneet etsiä tiedon lähdettä ja ihmistä, jolta voisi kysyä uudempaa tietoa. Tämä saman suuntainen ilmiö kuvattiin siis myös poliisin toiminnassa Ruotsissa tehdyssä tutkimuksessa (Borglund 2005).

Vaikka Työperäisten sairauksien rekisteristä tietoja pyytävät henkilöt eivät ole etsimässä vain jotakin epämääräistä terveyteen tai terveystietoon liittyvää tietoa, on silti mielenkiintoista, että tutkijat ovat jo muutama vuosi sitten huomanneet kasvavan trendin terveystietojen etsimisessä verkosta. Tämä johtaa väistämättä jatkossa terveysasioihin liittyvän tietokäyttötymisen tutkimuksen kasvuun, ja se on laitettu merkille jo muun muassa USA:ssa, Kanadassa, Suomessa ja Australiassa. (Case & Given 2016, 365.)

### **3.3 Rekisteritutkimus**

Tässä luvussa käydään läpi rekistereihin liittyviä tutkimuksia. Tarkoitus on selvittää, millaisia näkökulmia rekisteritutkimuksista löytyy, ja mitä johtopäätöksiä tutkimuksissa on saatu aikaiseksi. Tätä tutkimusta vastaavia rekisteritutkimuksia ei juurikaan ole tehty. Rekistereiden sisältöä on maailmalla ja Suomessa tutkittu substanssia eli rekistereiden

sisältöä analysoimalla, mutta itse rekistereiden tietosisällön ja tiedon loppukäyttäjien analysoinnista on hyvin vähän kirjallisuutta.

Artikkeli “Finnish health and social welfare registers in epidemiological research” (Gissler & Haukka 2004) kertoo Suomen lakiin pohjautuvista terveydenhuollon ja sosiaalisen hyvinvoinnin rekistereistä. Kyseessä on oikeastaan ainoa lähde, johon voi nojata tässä työssä tarkoitettujen rekisteritutkimuksen osalta. Tämäkään ei ole aiheen varsinaisessa merkityksessä tietyn rekisterin tutkimus, vaan kartoitus, millaisia Suomen terveydenhuollon ja sosiaalisen hyvinvoinnin rekisterit ovat, ja mikä on niiden historiallinen tausta. Artikkelissa on analysoitu myös niiden datan käyttöä.

Toinen aiheeseen liittyvä artikkeli on Iranissa tehty tutkimus, jossa on arvioitu vastaavalaisten työperäisten sairauksien rekistereiden mittakaavaa kymmenessä maassa ympäri maailman niiden käytön, sisällön, ja eräiden muiden tärkeiden asioiden kautta. Suomen TPSR on yhtenä rekisterinä otettu mukaan tutkimukseen. Tutkimus on kuitenkin suppea, koska se on lähinnä listaus kymmenen maan rekisterin minimimuuttujista, datalähteistä ja luokitusjärjestelmistä. Artikkeli perustuu kirjallisuuteen, ja Suomen tiedot on otettu mukaan TPSR:n englanninkielisistä vuosijulkaisuista vuosilta 2002 ja 2012 (Davoodi ym. 2017).

Henkilötunnus otettiin Suomessa käyttöön vuonna 1964. Tämä johti siihen, että lähes kaikki hallinnolliset rekisterit alkoivat käyttää sitä siitä lähtien. Vaikka tietosuojalait ovat Suomessa tiukat, voidaan henkilötietoja sisältävää dataa käyttää tieteellisiin tutkimuksiin. Tällaiset rekisterit ovat arvokkaita, koska pyrkimys on kansalaisten terveyden, terveyspalvelujen ja ylipäätään hyvinvoinnin parantamiseen. Tärkeiden arkistojen pito Suomessa on perinteikästä, ja niitä on ylläpidetty jo vuosikymmeniä. Epidemiologisten rekistereiden tutkimuksia on alettu tehdä enemmän ja tutkimuksista suoriudutaan helpommin, kun

tietotekniikan kehittyessä rekisterit ovat siirtyneet sähköiseen muotoon. (Gissler & Haukka 2004, 113.)

Vuonna 1987 Suomessa tuli voimaan laki, joka koski nimenomaan terveystietojen rekistereitä. Laki takasi kansalaiselle muun muassa oikeuden yksityisyyteen, vaikka rekistereihin oli alettu kerätä arkaluontoisia tietoja. Laki myös nosti esille tarpeen kerätä tätä arkaluontoista dataa lääketieteen ja terveystietojen osalta. Laki myös listasi kaikki ne rekisterit, joita valtakunnalliset viranomaiset saivat kerätä ja hallita. Näille rekistereille on yhteistä se, että niistä löytyvä tieto ei voi koskaan olla osa yksilön terveydentilaa koskevaa päätöksentekoa, vaan tieto on enemmänkin tilastotarkoituksiin kerättyä. Yksilön terveydentilaa koskevissa päätöksissä ovat edelleen hoitavat tahot ja niiden ylläpitämät potilastietojärjestelmät tärkeimpiä ja ainoita lähteitä. (Gissler & Haukka 2004, 115.)

Terveydenhuollon tietokantoja käytetään enenevässä määrin epidemiologisiin tutkimuksiin. Niiden tarkoituksena on muun muassa tutkia lääketieteellisten tuotteiden turvallisuutta, tehokkuutta ja käyttöä terveydenhuollossa. Tällaisissa tietokannoissa ja tutkimuksissa on kuitenkin pulmana se, että niissä on paljon sekaannuksia aiheuttavia tekijöitä. Sekaannuksia ja vääriä tulkintoja voivat aiheuttaa erilaiset muuttujat, potilaiden tausta, lääkärin osaaminen ja erityisala, sekä puutteelliset ja vain valikoiden tallennetut tiedot sähköisiin potilastietojärjestelmiin. (Brookhart ym. 2010, 114-116.) TPSR:n datan laadussa huomattuja pulmia ei kutsuta sekaannuksiksi, vaan puhutaan yleisesti datan laadusta. Kun kyseessä on data, joka kulkee monen eri toimijan ja käsittelijän kautta, ja matkan varrella dataa muun muassa harmonisoidaan eli yhdenmukaistetaan, on todennäköistä, että alkuperäinen tieto ei ole täysin samanlaista enää silloin, kun se päättyy rekisteriin (esim. Koskela ym. 2021, 78-79).

Jotta rekistereiden käyttö olisi järkevää, niiden datan laadun täytyy olla luotettavaa. Datan täytyy olla kattavaa ja laadullisesti oikeaa (Gissler & Haukka 2004, 115). Työperäisten sairauksien rekisterissä datan hyvä laatu on yksi tärkeimpiä ominaisuuksia. Viimeisimpänä tutkimuksena, joka TPSR:stä on tehty, oli tuoreimpien kymmenen rekisteröintivuoden datan analysointi, jonka tarkoituksena oli muun muassa löytää riskiammatteja, -aloja ja -altisteita (Koskela ym. 2019). Samalla kun dataa kerättiin ja analysoitiin, tehtiin myös laaja laadullinen tarkistus, jotta varmistuttiin että data vastaa uusimpien tilastomäärittysten ja vaatimusten mukaisia standardeja.

Vuonna 2002 Gissler ja Haukka ovat analysoineet tutkimuksia, joita on tehty tietyille suomalaisille rekistereille keskittyen niiden sisällöllisen oikeellisuuden tarkistamiseen. Aiemmin on tehty useita tutkimuksia myös siten, että on verrattu potilaan alkuperäisiä primäärilähteestä saatuja tietoja rekisterin sisältämiin tietoihin. Tutkimukset ovat osoittaneet tutkimuksen kohteena olleet rekisterit laadukkaiksi. (Gissler & Haukka 2004, 115.) Työperäisten sairauksien rekisteriin tuleva tieto ei välttämättä ole täysin samanlaista kuin mitä se on primäärilähteessä, eli potilastietojärjestelmässä. Primäärilähteenä voidaan pitää myös vakuutusyhtiön kirjaamia tietoja, ja myös ne saattavat muuttua alun ammattitautiepäilystä (ensimmäinen diagnoosi) vahvistettuun ammattitautiin (lopullinen diagnoosi). Potilaan tiedot ovat ennen TPSR:ään päätyä kulkeneet siis monen eri järjestelmän ja toimijan kautta: hoitavan lääkärin, vakuutusyhtiön, sekä Tapaturmavakuutuskeskuksen järjestelmien ja tietoja käsittelevien ihmisten kautta (ks. esim. kuva 1). Siksi tietojen tarkastaminen ja laadullinen korjaaminen on rekisterin vastuuhenkilöillä tärkeä työvaihe.

### 3.3.1 Laadukas rekisteri

Rekisteritutkimuksissa on käynyt ilmi, että rekistereiden data on sitä laadukkaampaa, mitä paremmin yhteistyö sujuu rekisterin hallinnoijan ja tietojen raportoivan tahon välillä. Myös rekisterin tietojen aktiivisen käytön, varsinkin päätöksenteossa ja tutkimuksissa, on huomattu parantavan laatua. (Gissler & Haukka 2004, 115.) Työperäisten sairauksien rekisterin vastuhenkilöt tekevät yhteistyötä virallisten tilastonpitäjien eli Tapaturmavakuutuskeskuksen ja Maatalousyrittäjien eläkelaitoksen kanssa, jotta datan laatu pysyisi mahdollisimman hyvänä. Tapaturmavakuutuskeskus ja sen alaisuudessa toimiva Tapaturma-asiain korvauslautakunta (TAKO) puolestaan ohjaavat ja neuvovat yleisesti ammattitaidon ilmoittamisessa sekä määrittelevät muun muassa vakuutusyhtiöiden kirjaa- ja korvauskäytäntöjä (Tapaturmavakuutuskeskus 2021a ja 2021b).

Rekisterissä olevaa tietoa käytetään monissa tutkimuksissa ja myös päätöksenteossa, kun havaitaan esimerkiksi jonkin sairauden tai altisteen yleistymistä. Tilastotiedon avulla voidaan kohdistaa erilaisia ennaltaehkäisy- ja tiedotuskampanjoita esimerkiksi riskitoimialoille tai -ammateille, joissa huomataan tietynlaisen altistumisen kasvua. (Koskela ym. 2019, 7, Laiho 2020.)

Rekisteritutkimus on kohdannut joitakin vastoinkäymisiä, kun on tutkittu rekistereiden käytön erilaisia mahdollisuuksia. Rekisteritiedon käyttöä hidastaa muun muassa lainsäädäntö, joka saattaa olla tiukkaa tai liian monimutkaista. Toisena pulmana on se, että terveys- ja sosiaalipuolen rekistereistä ei välttämättä tiedetä kovinkaan paljoa, tai että mitä ne sisältävät. Kolmanneksi ongelmaksi on nostettu tutkijoiden ja opiskelijoiden huonot metodologiset taidot rekisteritutkimuksen osalta. On myös kritisoitu mahdollista korkeaa hintaa, joka saattaa jopa estää datan hankkimisen tiettyjen toimijoiden rekistereistä tai

tietokannoista. Jatkuvasti tiukkenevat tietoturvalait koetaan kuitenkin pahimmaksi uhaksi rekistereiden käytölle tutkimustarkoituksiin. (Gissler & Haukka 2004, 118-119.)

Työperäisten sairauksien rekistereitä tutkimalla on saatu selville, että ammattitaudit lisääntyvät maailmanlaajuisesti. Vanhoja altistavia tekijöitä poistuu, mutta samalla uusia altisteita ja myös altistavia ammatteja nousee esiin. Monesti jo tapahtunutta altistumista ei voida parantaa, mutta se voitaisiin todennäköisesti ennaltaehkäistä. Siispä jokaisessa maassa tulisi rekisterin yhtenä tärkeänä tavoitteena olla ammattitautien ehkäisy ja mahdollisimman aikainen havaitseminen. Tutkimuksissa on korostettu myös sitä, että rekistereiden tunnuspiirteiden ja muuttujien tulisi vastata toisiaan maittain, jotta dataa voitaisiin verrata maailmanlaajuisesti. (Davoodi ym. 2017.) Eurostatilla on meneillään kaikkia EU-maita sekä muutamia EU:n ulkopuolisia Euroopan maita koskeva pilottihanke, jossa kehitetään ammattitautien tilastointia ja eri maiden ammattitautitilastojen vertailtavuutta (EUR-Lex 2021).

## 4 TUTKIMUSASETELMA

Luvussa 4.1 esitellään tutkimuskysymykset ja tutkimusmenetelmä. Luvussa 4.2 selvitetään tarkemmin, millaisesta datasta rekisterin tietosisältö muodostuu. Se toimii myös pohjana ymmärtää tutkimuksen toteutusta. Luku 4.3 keskittyy tutkimuksen toteutukseen.

### 4.1 Tutkimuskysymykset ja tutkimusmenetelmä

Tutkimuksessa selvitettiin Työperäisten sairauksien rekisterin sisältämän tiedon käyttöä analysoimalla rekisteriin tulleet tietopalvelupyynnöt. Tutkimuksen tavoitteena oli saada selville ilmiöitä rekisterin käytöstä, jotta jatkossa voidaan paremmin kehittää rekisterin julkaisuja. Selvitetiin, kuinka paljon tietoa pyydetään, mihin tarkoituksiin sitä käytetään, kuka tietoa käyttää, ja millaisia hakuheitoja pyynnöillä oli.

Tutkimuskysymykset ovat:

1. Millainen tausta tiedon pyytäjillä on ja mihin tarkoitukseen tietoja käytetään?
2. Millaisia tietopalvelupyynnöitä rekisteriin on tehty neljän vuoden aikana, ja kuinka paljon?

Ensimmäisen tutkimuskysymyksen tavoitteena oli selvittää, millainen on tietopalvelupyynnöiden tekijöiden tausta, ja mitä tarkoitusta varten pyyntö on tehty. Toisen tutkimuskysymyksen tarkoituksena oli selvittää tietopalvelupyynnöiden sisältöä, eli millaista dataa ja kuinka paljon sitä on rekisteristä pyydetty.

Tähän tutkielmaan suunniteltiin alun perin tehtäväksi myös kyselytutkimus, joka olisi ollut tietopalvelupyynnöiden tekijöille suunnattu. Kysely päätettiin kuitenkin jättää työstä pois siihen liittyneiden tietosuojasioiden sekä julkisuuslakiin liittyvien kysymysten

vuoksi. Varsinainen tutkimus keskittyi siis tietopalvelupyynnöistä koostuneen aineiston analysointiin.

Tällä pro gradu -tutkimuksella pyrittiin tuottamaan yksityiskohtaista tietoa rekisterin käytöstä analysoimalla tietopalvelupyyntöjen sisältöä, jotta aiheeseen saadaan uusia näkökulmia ja vähän tunnetut ilmiöt saadaan esille. Siispä sen voi kuvailla olevan *kartoittava analyysi* rajattuun aiheeseen kuuluvasta datasta, joka vastaa kysymyksiin kuka, mitä ja kuinka paljon. (Vilkkä 2007, 20, Hirsjärvi ym. 2015, 138.) Tutkimus on samankaltainen, mutta ei luonteeltaan kuitenkaan täysin samanlainen kuin esimerkiksi *tapaustutkimus*, joka tuottaa yksityiskohtaista tutkimustietoa yksittäisestä tapauksesta. Tapaustutkimukseen ei ole tekniikka, jolla aineistoa kerätään, eikä se määrittele mitä menetelmiä tutkimuksessa käytetään, mutta sen tarkoituksena on löytää vastaus kysymyksiin, kuinka ja miksi. (Valli 2015, 180-183.)

Kvantitatiivisessa eli määrällisessä tutkimuksessa on olennaista, että aineistoa voidaan mitata ja käsitellä siten, että tulokset voidaan kuvailla numeerisessa muodossa. (Vilkkä 2007, 27-31, 35, Hirsjärvi ym. 2015, 136-137.) Kvantitatiivisen tutkimuksen aineiston kerääminen eli havaintomatriisin muodostaminen, käsittely ja analysointi ovat kaikki eri vaiheita. Aineisto käsitellään tutkimusongelman ratkaisu mielessä, eli siten, että tutkimuskysymyksiin saadaan vastaus. (Heikkilä 2014, 120, 138.) Jotta analyysivaiheessa päästään virheettömään lopputulokseen tulkintojen kannalta, on mitattavien arvojen tallennus syytä tarkistaa ajoissa, eli muuttujien määrittely niiden nimien ja arvojen osalta (Vilkkä 2007, 114). Analyysivaihe on tärkeä, koska silloin selviää, onko tutkimusongelmat asetettu oikein, ja voidaanko tuloksia tulkita niin, että niistä on mahdollista tehdä johtopäätöksiä (Hirsjärvi ym. 2015, 221, 229-230).



## 4.2 Rekisterin tietosisältö

Tutkimuksen toteutuksen taustaksi ja ymmärtämiseksi on olennaista selvittää Työperäisten sairauksien rekisterin tietosisältöä. Rekisteri sisältää useita erityyppisiä muuttujia. Rekisterin alkuaikojen datassa tapausten tiedoissa on ollut vähemmän muuttujia kuin nykymuotoisten tietojärjestelmien aikana. Muuttujia on vuosien aikana lisätty järjestelmien kehityksen myötä, sekä siksi, että on tunnistettu uudenlaisia tapoja ja tietoja, joita potilastiedoissa ja sen myötä rekisteröinnissä on haluttu tallentaa.

Muuttujista rekisteröintivuosi, laji (onko tapaus ammattitauti vai ammattitautiepäily) ja mahdollisen ammattitaudin vahvistamispäivä löytyy jokaisesta tapauksesta, joka on rekisteröity vuoden 2003 jälkeen. Henkilön eli vahingoittuneen tiedoista henkilötunnus, ammatti ja sairauden diagnoosi ovat tärkeitä tietoja. Ammattitautiepäilytapauksille laskeaan ikä henkilötunnuksen ja rekisteröintivuoden perusteella, ja vahvistetuille ammattitautitapauksille henkilötunnuksen ja vahvistamispäivän perusteella. Henkilötunnuksen perusteella määritellään myös tapauksen sukupuoli. Tällä hetkellä käytössä on ICD10-diagnoosikoodisto (Terveystieteiden tutkimuskeskus 2011) sairauksille.

Tapauksilla on Tilastokeskuksen virallisten ammatti- ja toimialakoodien mukainen ammatti- ja toimialatieto. Koodistot ovat muuttuneet vuosien saatossa useita kertoja, ja vanhoja ja uusia koodeja pyritään käyttämään yhteisissä tilastoissa koodien luokitusavainten avulla. Uusimmat ovat TOL2008 ja AML2010 -luokitukset. (Tilastokeskus 2021.) Toimiala määräytyy työnantajan toimialan mukaan, kuten myös tieto (paikka)kunnasta määräytyy työnantajan mukaan. Kuitenkin moni työnantajan mukaan määräytyvä toimiala antaa väärän kuvan työn luonteesta, joten ne korjataan rekisterissä vastaamaan työtä, jota oikeasti tehdään. Tästä esimerkkinä on toimiala 84 Julkinen hallinto ja maanpuolustus;

pakollinen sosiaalivakuutus. Sen alle ilmoitetaan esimerkiksi kaikki kunnille kuuluvat alat, kuten koulutus, sekä terveydenhuollon ja sosiaalialan työt. Jotta saadaan oikea kuva, paljonko kullakin toimialalla oikeasti sairastutaan ammattitauteihin, pitää tietoa korjata tapauksen muiden tietojen, kuten ammatin perusteella.

Vahingoittuneen altistumistieto on tärkeä muuttuja. Rekisteriin tuleva data sisältää TVK:n ja Melan omien altistekoodien mukaisen melko yleisluontoisen tiedon altisteesta. Datan mukana tulevan sanallisen ”kuvaus”-tiedon tulisi tilastoimisohjeen mukaan pitää sisällään mahdollisimman tarkka tieto altisteesta, mutta monissa tapauksissa tämä tieto puuttuu kokonaan. Rekisterissä tapauksille päivitetään altistekoodi Työterveyslaitoksen oman Systemaattisen altistekoodiston mukaan, joka sisältää noin 3000 erilaista kuusi-merkkistä altistetta. Näistä valitaan oikea koodi jokaiselle tapaukselle. Altisteluettelosta löytyy esimerkiksi kemiallisia, biologisia ja fysikaalisia altisteita.

Rekisteriin tulevan datan lisäksi tutkitaan aluehallintovirastoista saatu lääkäreiden tekemä ilmoitus ammattitaudista tai ammattitautiepäilystä. Joskus ilmoituksen avulla saadaan tapauksesta tärkeää täydentävää lisätietoa, mutta toisinaan ilmoituskaan ei anna lisää informaatiota. Ammattitauti-ilmoitus löytyy vain pieneltä osalta rekisterin tapauksista, noin neljäsosalta kokonaismäärästä vuosittain (Koskela ym. 2021, 8-9).

Melko uusina muuttujina ovat tautikohtaiset altisteryhmät. Ryhmittely on luotu sekä tutkijoiden että käyttäjien tiedontarpeiden pohjalta ja sen tekemiseen ovat osallistuneet muun muassa kemisti, iho-, keuhko- sekä korva-, nenä- ja kurkkusairauksien erikoislääkäreitä. Ryhmittelyn avulla pystytään helpommin poimimaan rekisteristä tilastoja tietyn tyyppisistä tapauksista.

Rekisterin tietosisältö kattaa myös monia erilaisia vakuutusyhtiön määrittelemiä muuttujia, kuten vahinkonumeron, yhtiötunnuksen ja sattumisvuoden. Vahingon seuraamustieto ilmaisee, kuinka vakavasta tapauksesta on kyse. Seuraamustietoja on erityyppisiä liittyen siihen, onko kyseessä Euroopan tilastointijärjestelmän yhteinen EODS-seuraamustieto, vai Suomessa käytössä oleva seuraamustieto. Yhtenä ammattitaudin seuraamustietona voi olla henkilön kuolema, jos kuolema johtuu ammattitaudista.

### **4.3 Tutkimuksen toteutus**

Tutkimuksen aineisto kerättiin neljän vuoden aikana tulleista tietopalvelupyynnöistä, joita TPSR:n vastuuhenkilöt olivat saaneet sähköpostin tai puhelimen välityksellä. Rekisterin sisältämää varsinaista potilasdataa ei tässä tutkimuksessa otettu huomioon. Tiedot, joita kerättiin, olivat palvelupyyntöjen tekijöiden taustatiedot yleisesti, sekä pyyntöjen sisällöt liittyen siihen, millaisia muuttujia niissä oli käytetty rekisteritiedon hakemiseen.

Aineistosta poistettiin yksi tapaus, joka liittyi suomalaiseen hankkeeseen ”Työkyvyn tila Suomessa”. Kyseinen tietopalvelupyyntö ei kohdistunut rekisterin varsinaisiin muuttujiin, vaan rekisteristä etsittiin henkilötunnusten avulla tietoja, jotka toimitettiin hankkeen vastuutaholle, joka yhdisti nämä tiedot muista rekistereistä saatuihin tietoihin. Hankkeen populaatio oli yli miljoona tapausta, eli kyseessä oli Suomen mittakaavassa laaja tutkimus. Tämän jälkeen aineistoon jäi 94 tietopalvelupyyntöä.

Aineiston keräämisen ajankohta oli marraskuun 2016 ja lokakuun 2020 välinen aika. Pyynnöt koottiin yhdeksi aineistoksi eli havaintomatriisiksi taulukko-ohjelmaan (Excel), josta käy ilmi pyynnön ajankohta, pyytäjä ja kuvaus millaista tietoa rekisteristä etsittiin. Tietopalvelupyynnöt olivat alun perin sanallisessa muodossa, joten kaikki havainnot muunnettiin ensin numero- ja merkkimuotoisiksi. Yhdellä rivillä oli yksi

tietopalvelupyyntö. Pyynnön sisällöstä sekä pyytäjän tiedoista luotiin rekisterin tietosivallön mukaisia muuttujia pystysarakkeiksi. Muuttujia olivat pyyntövuosi, pyytäjän sukupuoli, pyytäjän tausta ammatillisessa mielessä, pyynnön tausta tiedon käyttötarkoituksen perusteella, sekä varsinaiset muuttujat (taulukko 1) tietopalvelupyynnön hakuehdoista.

Pyynnöistä palasteltiin omiksi muuttujikseen pyytäjä sekä pyynnön kohteena oleva data. Pyynnön tekijän henkilötiedot (nimi ja sähköpostiosoite) poistettiin aineistosta. Pyynnön tekijät luokiteltiin 1) opiskelija, 2) erikoistuva lääkäri ja 3) työssäkäyvä. Pyytäjän työnantajaa tai opiskelupaikkaa ei myöskään otettu mukaan tutkimukseen. Myös pyynnön tarkoitus luokiteltiin kolmeen ryhmään. Nämä olivat 1) opinnäytetyö/syventävä työ, 2) asiantuntijakäyttö/tutkimushanke ja 3) julkaisu (lehti, kirja yms.). Pyytäjän sukupuoli määriteltiin ainoastaan nimen mukaan: 1=mies, 2=nainen ja tässä muuttuja sai samat arvot kuin rekisterissä, jossa tapauksen sukupuoli määritellään kuitenkin henkilötunnuksen mukaan.

Edellä mainittuihin luokituksiin sisällytettiin siis taustan mukaan kaikki pyynnöt. Yhdeksi ryhmäksi luokiteltiin johonkin oppilaitokseen (yleensä korkeakoulu) lopputyötä tekevät henkilöt eli opiskelijat, eli heidän pyyntönsä tarkoitus liittyy opinnäytetyöhön. Toisena ryhmänä opintoihin liittyviä tietopyyntöjä tekivät työterveyshuoltoon erikoistuvat lääkärit. Heidän erikoistumisopintoihinsa kuuluu muun muassa kuuden kuukauden erikoistumisjakso Työterveyslaitoksella, ja siihen kuuluu lopputyön (syventävä työ) tekeminen. Osa erikoistuvista lääkäreistä käyttää lopputyössään ammattitautitilastoja.

Kolmanteen tässä työssä määriteltyyn tietopalvelupyyntöjä tehneeseen käyttäjäryhmään lisättiin kaikki työssäkäyvät henkilöt, eli he, jotka eivät pyydä tilastoja opintoihin liittyviin asioihin. Näitä ovat esimerkiksi tutkijat (erityisesti Työterveyslaitokselta), lääkärit

esimerkiksi yliopistosairaaloista, ammattijärjestöjen ihmiset, jotka ovat kiinnostuneita ammattitaukeista liittyen edustamiinsa toimialoihin ja ammatteihin. Myös vakuutuslaitokset, Sosiaali- ja terveysministeriö sekä aluehallintovirastot, sekä jotkut yksityiset toimijat ovat kiinnostuneita tarkemmista tilastoista, joita ei saa suoraan selville TPSR:n vuosijulkaisusta. Joskus yksityinen ammatinharjoittaja saattaa pyytää tilastoja, kun hän luennoi tietyistä sairauksista jollekin ammattiryhmälle, ja haluaa nostaa esiin juuri sitä ryhmää koskevat ammattitaudit. Nämä kaikki luokiteltiin pyynnön tarkoituksen mukaan asiantuntijakäyttöön kuuluvaan ryhmään.

Joillakin toimialoilla tai ammatillisissa yhdistyksissä tehdään omia vuosittaisia kirjoja, lehtiä tai muita julkaisuja, ja näihin pyydetään tilastotietoja TPSR:stä. Tämän kaltaiset pyynnöt luokiteltiin tarkoituksen mukaan ryhmään julkaisu. Esimerkiksi Ruokatieto Yhdistys ry julkaisee vuosittain ”Tietohaarukka – tilastotietoa elintarvikealasta” -tietopakettin, jossa on mukana myös alaan liittyvien ammattitautien tilastoja (Ruokatieto 2021). Myös media on kiinnostunut ammattitaukeista. Vuosijulkaisuja referoidaan ja ammattitaukeista julkaistaan artikkeleita. Mediasta otetaan yhteyttä rekisterin vastuulääkäriin myös haastattelujen merkeissä, esimerkiksi asbesti- ja kosteusvaurioasioissa, kun halutaan tietoa näihin liittyvistä ammattitaukeista. Mediaan liittyvät pyynnöt luokiteltiin tässä työssä asiantuntijakäytön ryhmään.

Varsinaisella tietopalvelupyynnöllä on aina melko tarkat hakuehdot. Pynnön tekijän mielenkiinto koskee monesti jotain tiettyä toimialaa tai ammattiryhmää, joten nämä otettiin mukaan erillisinä muuttujina, aivan kuten ne ovat rekisterissäkin. Rekisterissä ammatti- ja toimialakoodi voivat olla määriteltynä hierarkkisina tasoina 1–5 merkkisenä, joista 5-merkkinen on tarkin (Tilastokeskus 2021). Tässä työssä käytettiin kahden merkin

tarkkuutta, jota käytetään myös rekisterin vuosijulkaisuissa, tai muuta tarkkuutta, jolla voitiin eritellä ammatit ja toimialat mahdollisimman informatiivisesti.

Usein jokin tietty sairaus tai altiste voivat olla pyynnön ehtoina, joten myös ne hyväksyttiin tutkimuksessa omiksi muuttujikseen. Tietopalvelupyynnöjen mukaan rekisteriin tehtävissä hauissa käytetään aina mahdollisimman tarkkoja diagnoosi- ja altistekoodoja. Varsinaisia koodeja voi tietokantaan tehtävissä sql-hakulausekkeissa olla yhdestä moneen kymmeneen. Tässä työssä ei ollut tarkoituksenmukaista selvittää eikä analysoida tarkkoja koodeja, ainoastaan tutkittiin ylätasolla mistä sairaudesta on haluttu tietoja. Erityyppisillä astmoilla voi olla monia erilaisia diagnoosikoodeja, kuten esimerkiksi J45.0, J45.1 tai J45.9. Tässä työssä aineiston tarkat diagnoosikoodit muunnettiin vastaamaan tautiryhmiä hyvin yleisellä tasolla. Mikä tahansa näistä koodeista sai tässä tutkimuksessa tiedon ”astma”. Sama periaate koski kaikkia diagnooseja. Sairauksien tarkat ICD10-koodit ryhmiteltiin siis rekisterissä olevien pääasiallisten ja yleisimpien ammattitautiryhmien mukaisesti. Koska Työterveyslaitoksen systemaattisesta altistelueluettelosta ja rekisteristä löytyy yli 3000 eri altistekoodia, ei niitä ollut järkevää esittää yksittäisinä koodeina ja niminä. Altisteet muunnettiin ensin yleisemmiksi altisteryhmiksi ja sen jälkeen systemaattisen altistelueluettelon mukaan ylimmän tason ryhmiksi.

Tietopalvelupyynnössä saattaa olla edellä mainittujen sijasta, tai niiden lisäksi, vielä jokin muu tarkentava tieto, esimerkiksi maakunta tai tapauksen ikä. Mahdollinen muu ehto sai tässä tutkimuksessa vain merkinnän, jos sellainen oli. Tilastopyynnöt eivät myöskään sijoitu mille tahansa aikamääreelle, vaan pyytäjän kiinnostuksen mukaisesti monessa tapauksessa tilastoja etsitään rekisteristä tiettyjen vuosien mukaan.

Taulukko 1. Tietopalvelupyynnöissä esiintyneet muuttujat.

Vuodet
Ammatti
Toimiala
Diagnoosi
Altiste
Muu hakuehto
Joidenkin näiden kombinaatiot

Esimerkiksi pyynnöstä ”sairaanhoitajien kosteusvaurioihin liittyvät astmat” löytyy ammattikoodi 32211 sairaanhoitajat, joka sai ryhmän sairaanhoitajat ym., diagnoosista astman koodi J45.9, joka kuuluu tautiryhmään *astma*, sekä altisteista kosteusvaurioon liittyvän homesienet-koodin, joka puolestaan kuuluu biologisiin altisteisiin.

Aineiston käsittelyn eli muuttujien tietojen koodaamisen ja ryhmittelyn jälkeen aineisto siirrettiin SPSS-tilastointiohjelmaan. Havaintomatriisi muuttuu tällöin järjestelmätiedostoksi, jota voidaan käsitellä muuttujakohtaisesti sekä luoda uusia muuttujia olemassa olevista tiedoista (Vilkka 2007, 113). Järjestelmätiedostoon luotiin uusi muuttuja *hakuehtojen määrä*, joka laski jokaisesta tietopalvelupyynnöstä eli yhdeltä riviltä lukumäärän, kuinka monta hakuehtoa eli eri muuttujaa pyynnössä oli käytetty.

SPSS-ohjelmassa aineisto analysoitiin ja tutkimustuloksia tuotettiin deskriptiivisen eli kuvailevan tilastotieteen menetelmillä. Tuloksissa esitettiin tilastollisia tunnuslukuja eli frekvenssejä sekä prosenttiosuuksia, joiden avulla pyrittiin tarkastelemaan rekisterin käyttöä. Näiden yksinkertaisten menetelmien avulla hahmotettiin yleiskuva pyynnön tekijöistä sekä pyyntöjen sisällöistä ja määristä. Ilmiöiden välisten yhteyksien selvittämiseen käytettiin kahden muuttujan ristiintaulukointia. (Alastalo & Borg 2010, Mellin 2006, Heikkilä 2014, 198.)

Tulokset esiteltiin ja havainnollistettiin taulukoiden ja kuvaajien avulla. Hyvä kuvaaja välittää oikean tiedon ja herättää lukijan mielenkiinnon yksinään ja riittävän hyvällä otsikoinnilla. Pylväskuvion avulla esitettiin havaintojen määriä, ja piirakkakuvion avulla kunkin muuttujan suhteellisia osuuksia aineistossa. (Vilka 2007, 138-143, Heikkilä 2014, 148-155.)

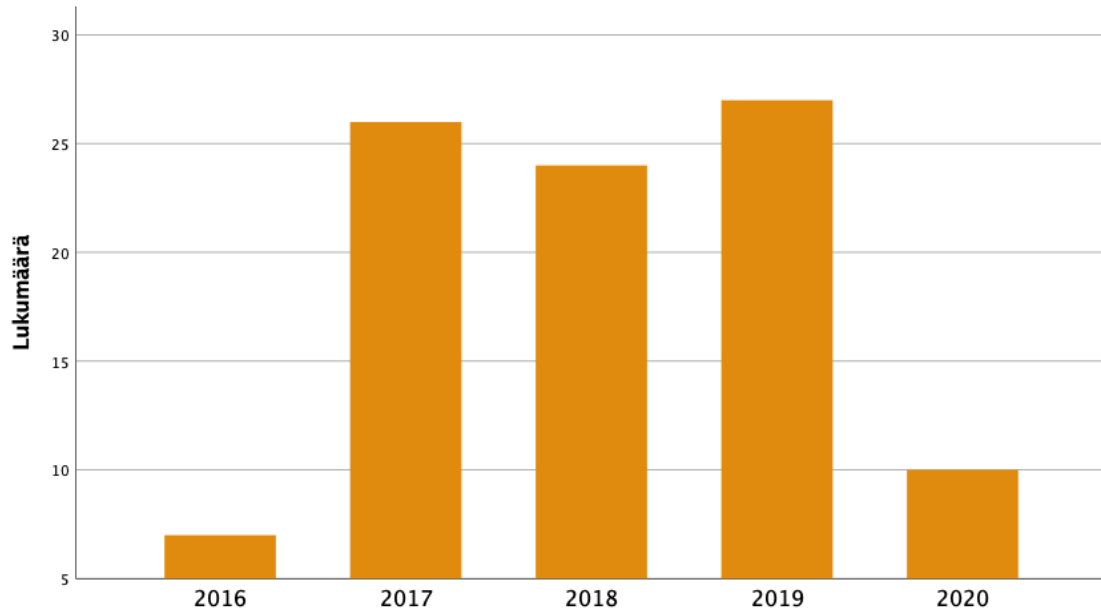


## 5 TULOKSET

Tässä luvussa esitellään tutkimuksen tulokset, ja ne on havainnollistettu taulukoiden ja kuvaajien avulla. Ensin kerrotaan tietopalvelupyyntöjen tekijöihin liittyvät asiat ja sen jälkeen varsinaisten pyyntöjen sisältämät havainnot muuttujista eli pyyntöjen sisältö. Lopuksi käydään läpi tulokset pyyntöjen sisällöistä käyttäjäryhmittäin tutkittuna.

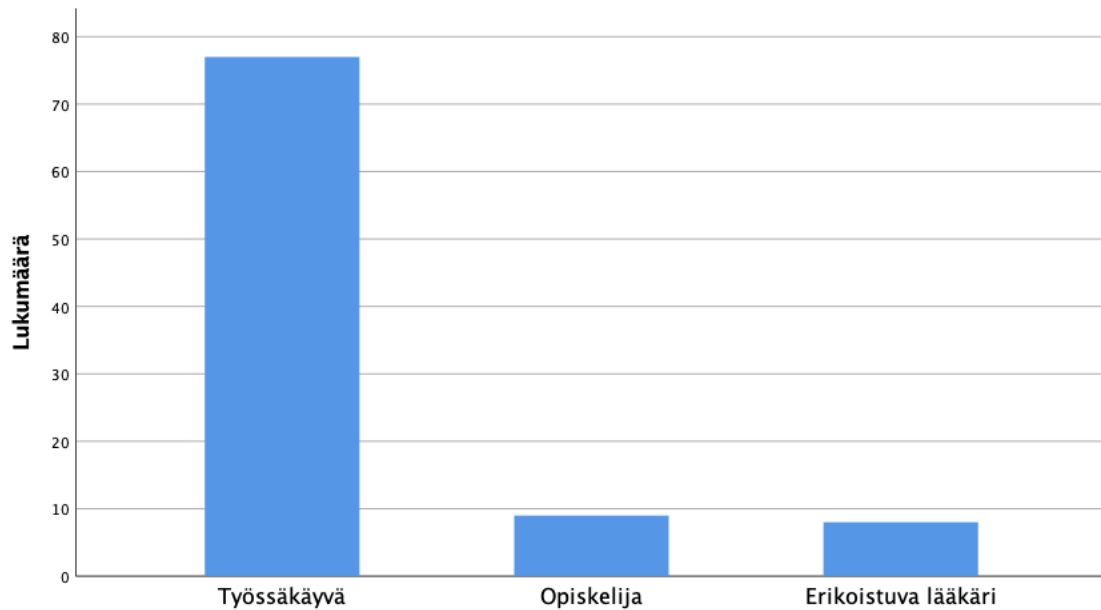
### 5.1 Tietopalvelupyyntöjen tekijät ja määrät

Tutkimuksessa oli tietopalvelupyyntöjä 94 kappaletta. Tietopalvelupyynnöillä saattoi olla sama pyytäjä useampaan kertaan, mutta jokainen pyyntö laskettiin yhdeksi itsenäiseksi tiedoksi. Kuvasta 2 nähdään, että tarkasteluaikana eniten pyyntöjä tehtiin vuonna 2019, 27 kappaletta. Koska vuosina 2016 ja 2020 ei kerätty tietoja koko kalenterivuoden ajalta, eivät näiden vuosien lukumäärät ole vertailukelpoisia muiden vuosien kanssa. Vuonna 2020 tietoja kerättiin kuitenkin suuri osa vuodesta (tammikuu-lokakuu), joten kuukausimäärään nähden pyyntöjen määrä 10 kappaletta, oli vain 37 % verrattuna vuoteen 2019. Pyyntöjen tekijöistä 69 oli naisia ja 25 miehiä.



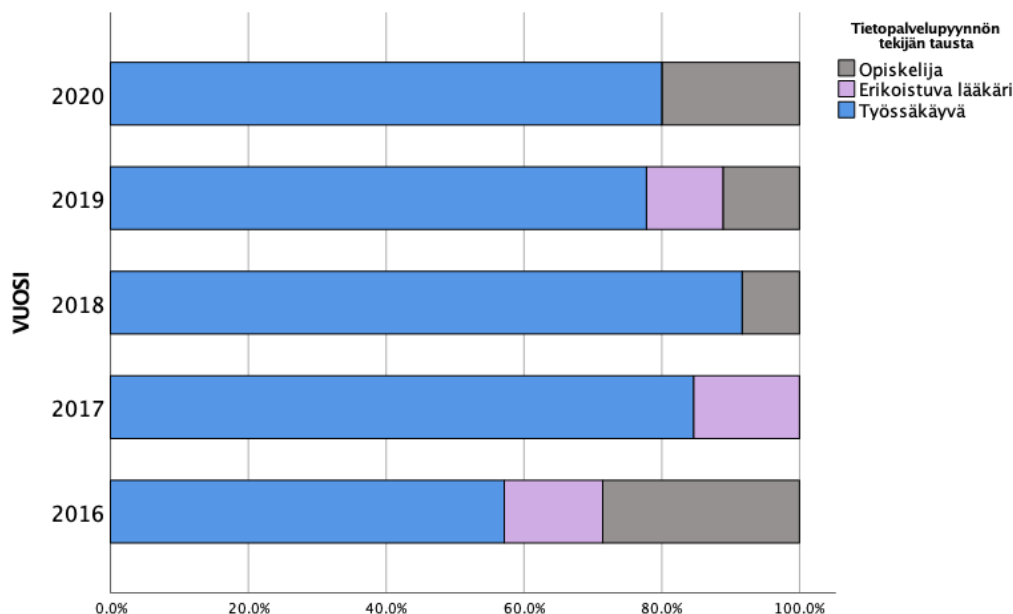
Kuva 2. Tietopalvelupyyntöjen määrät tarkasteluvuosittain.

Tietopalvelupyyntöjen tekijät jaettiin kolmeen ryhmään. Näistä ryhmistä selvästi suurin ryhmä oli *työssäkäyvä*, joka sisälsi kaikki muut paitsi ryhmään *opiskelija* tai ryhmään *erikoistuva lääkäri* kuuluvat henkilöt. Kuva 3 osoittaa, että ryhmään *työssäkäyvä* kuului 82 % kaikista pyyntöjen tekijöistä. Myös erikoistuvat lääkärit ovat *työssäkäyviä*, mutta koska he ovat samalla yliopisto-opiskelijoita, heidät jaettiin omaan ryhmäänsä.



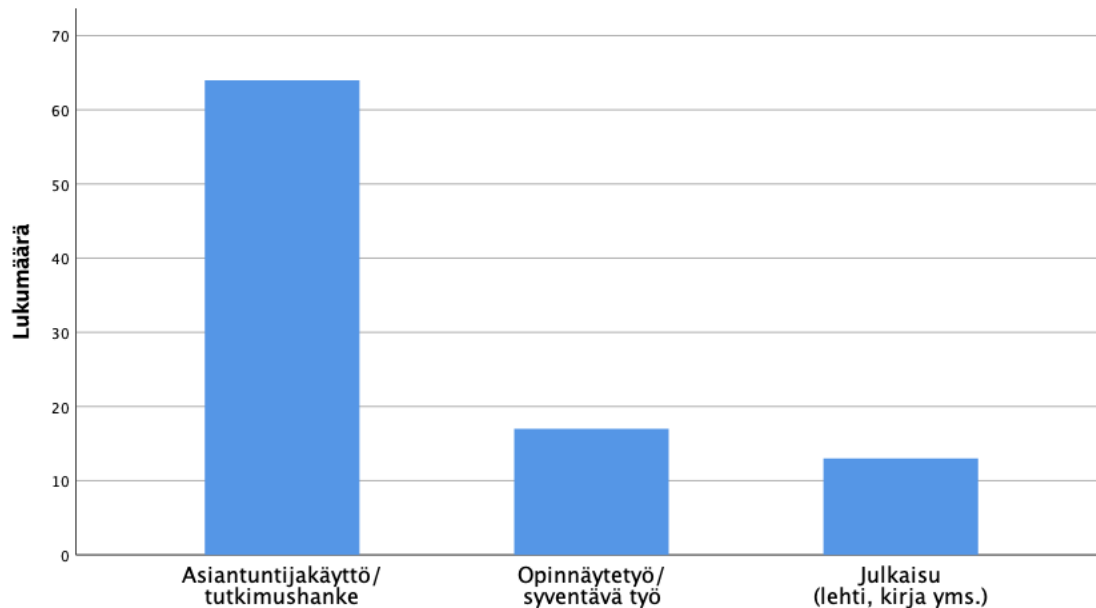
Kuva 3. Tietopalvelupyyntöjen määrät pyytäjän taustan perusteella.

Tuloksista kävi ilmi, että työssäkävien ryhmään kuuluvat pyytävät tietoja Työperäisten sairauksien rekisteristä melko tasaisesti vuosittain. Sen sijaan opiskelijoiden pyyntöjä ei tullut vuonna 2017 yhtään. Myöskään erikoistuvien lääkäreiden ryhmään kuuluvat eivät pyytäneet tietoja vuonna 2018 eivätkä vuonna 2020. Kuvassa 4 nähdään suhteelliset jakaumat vuosittain pyyntöjen tekijöiden mukaan, mutta on otettava huomioon, että vuodet 2016 ja 2020 eivät sisältäneet koko kalenterivuotta aineiston keräämisessä.



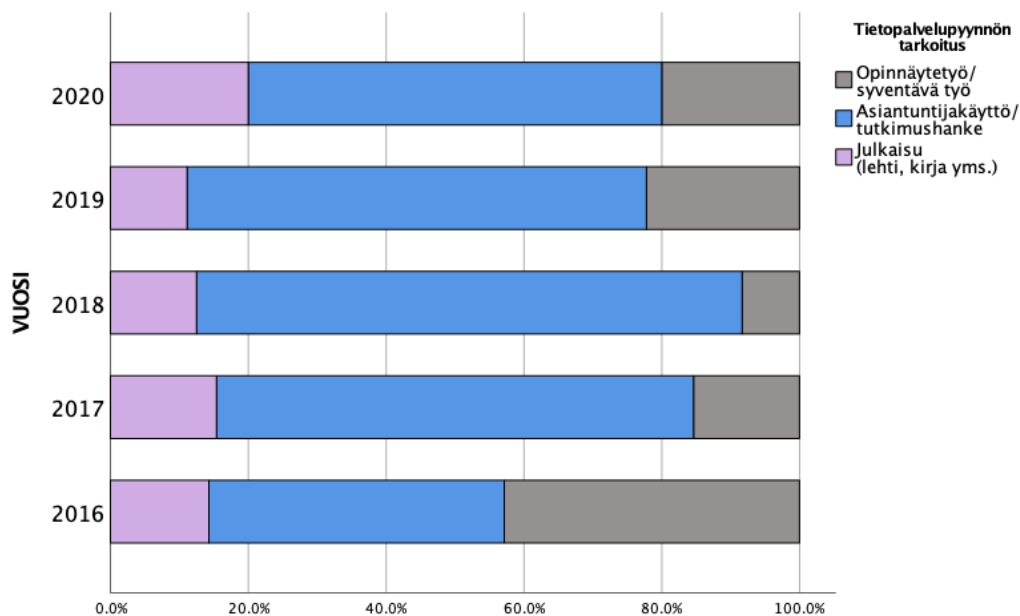
Kuva 4. Tietopalvelupyynnöiden tekijäryhmien suhteelliset osuudet vuosittain.

Yhtenä taustatietona oli tietopalvelupyynnön tarkoitus. Pyynnöt jaettiin niiden käyttötarkoituksen mukaan kolmeen eri ryhmään, jotta pystyttiin ilmoittamaan tulokset melko yleisellä tasolla. Kuvasta 5 käy selville, että suurimpana ryhmänä oli *asiantuntijakäyttöön tai tutkimushankkeeseen* pyydyt tiedot. Tähän samaan ryhmään kuuluvat myös kaikki mahdolliset median tekemät haastattelu- tai artikkelipyynnöt. Pääsääntöisesti nekin ovat asiantuntijakäyttöä. Asiantuntijakäytön osuus kaikista pyynnöistä oli 68 %. Muut ryhmät, eli *opintoihin* liittyvät pyynnöt (18 %) sekä *julkaisuja* (lehtiä, kirjoja ym. painettuja teoksia) varten tehdyt pyynnöt (14 %) olivat melko tasaisesti jakaantuneet.



Kuva 5. Tietopalvelupyyntöjen käyttötarkoitukset ja niiden määrät.

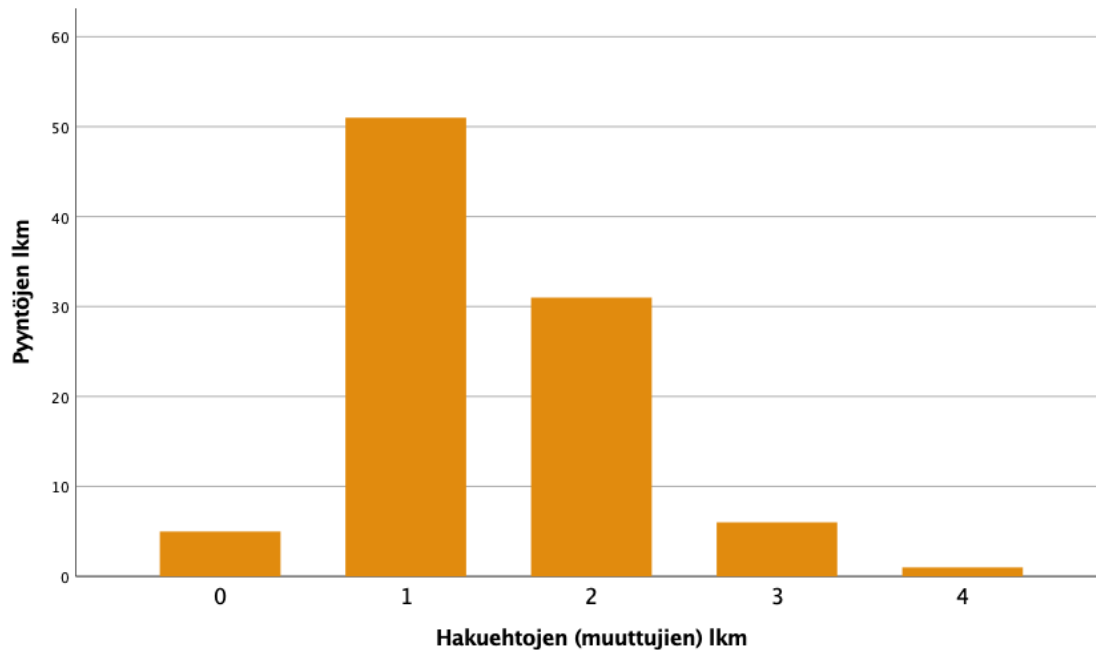
Kuvasta 6 nähdään tietopalvelupyyntöjen käyttötarkoitukset vuositasolla suhteessa toisiinsa. Asiantuntijakäyttöön tai tutkimushankkeeseen haluttuja tietoja on pyydetty melko tasaisesti joka vuosi. Selkeästi vähemmän tietoja pyydetään koulutukseen liittyviin opinnäyte- ym. töihin sekä painettuihin julkaisuihin, mutta melko tasaisesti niitäkin vuosittain on tullut. On kuitenkin muistettava, että monesti pyynnön tullessa rekisterin vastuuhenkilöille ei ole selvää, mihin tarkoitukseen tietoa lopulta käytetään. Joskus asiantuntijakäyttöön pyydettyä tietoa saatetaan käyttää esimerkiksi painetussa julkaisussa myöhemmin.



Kuva 6. Tietopalvelupyynnöiden tarkoitukset suhteellisina osuuksina vuosittain.

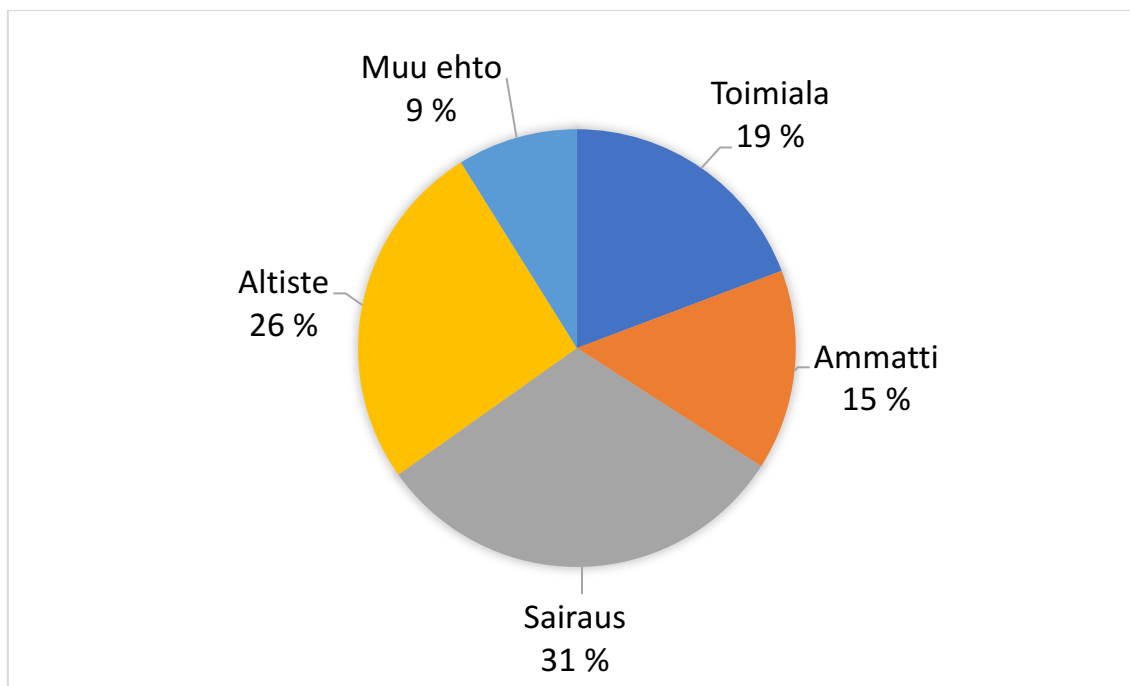
## 5.2 Tietopalvelupyynnöiden hakuehdot eli muuttujat

Tutkimuksen aineistoon luodun uuden muuttujan *hakuehtojen määrä* perusteella saatiin selville, että tietoa pyydetään rekisteristä useimmin yhden hakuehdon mukaan, kuten kuva 7 osoittaa. Tällaisia pyyntöjä oli 51 kpl, mikä oli 54 % kaikista pyynnöistä. Joskus hakuehtoja oli useampia kuin yksi samassa pyynnössä. Toiseksi suosituin oli kahden hakuehdon muodostama pyyntö, ja niitä oli 31 kpl, eli 33 % kokonaismäärästä. Myös kolmen ja neljän hakuehdon muodostamia pyyntöjä on tehty jokunen, mutta niitä oli yhteensä vain 8 % kaikista pyynnöistä. Sellaisia pyyntöjä, joissa ei ollut yhtään varsinaista hakuehdoksi määritellyttä muuttujaa oli 5 % kokonaismäärästä. Tarkentavan hakumuuttujan puuttuminen selittyy sillä, että kyseessä on pyyntö, joka kohdistuu esimerkiksi kaikkien dataan, joka löytyy tiettyjen vuosien aikana raportoiduista tapauksista rekisteriin. Eli silloin rajaavana tekijänä on ollut rekisteröintivuodet eikä varsinainen muuttuja.



Kuva 7. Hakuehtojen (muuttujien) lukumäärä per tietopalvelupyyntö.

Pyynnöissä olevat hakuehdot laskettiin jokainen yhdeksi muuttujaksi. Pyyntöjä oli yhteensä 94, mutta koska yhdessä pyynnössä saattoi olla monta muuttujaa, oli hakuehtojen yhteismäärä 135 kpl. Kuvasta 8 selviää, että yleisin hakuehto oli jokin sairaus tai tautiryhmä (42 kpl, 31 % kaikista muuttujista). Muita hakuehtoja olivat altiste (35 kpl, 26 %), toimiala (26 kpl, 19 %) ja ammatti (20 kpl, 15 %). Jokin muu hakuehto kuin aiemmat oli mukana toisen hakuehdon kanssa tai yksinään yhteensä 12 tapauksessa (9 % kaikista pyynnöistä). Muu hakuehto oli esimerkiksi kunta, maakunta, sairaanhoitopiiri, rekisteröidyn ihmisen ikä, ammattitautiin kuolleet tai henkilötunnus (kun kyseessä on ollut tutkimusluvan saanut hanke). Muun hakuehdon sisältävistä pyynnöistä kunta/maakunta/sairaanhoitopiiri oli haettavana tietona 33 %:ssa hauista.



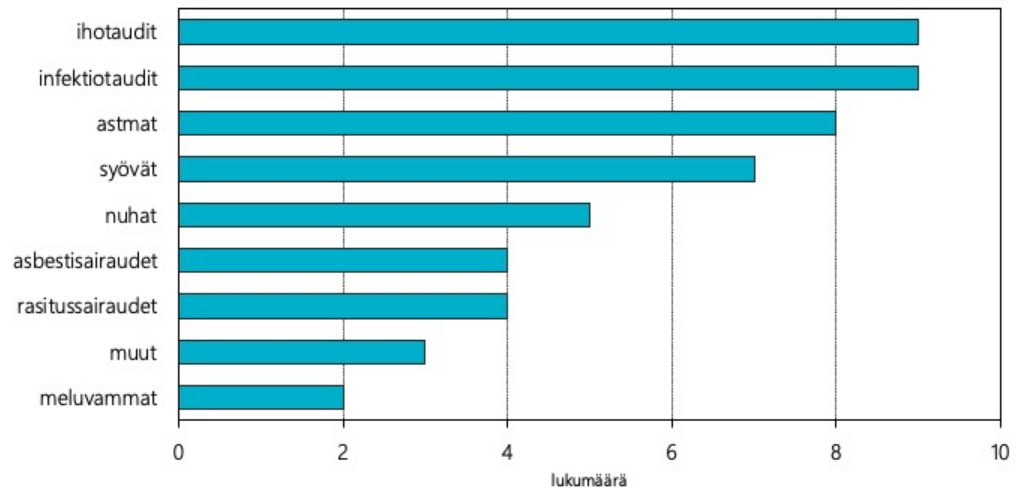
Kuva 8. Tietopalvelupyynnöissä esiintyneiden hakuehtojen eli muuttujien osuudet.

### 5.2.1 Tautiryhmä eli sairaus

Suosituin hakuehto oli jokin sairaus. Sairauden, eli yleensä ICD10-diagnoosikoodin mukaisia hakuja tehtiin 31 % kaikista pyynnöistä.

Astmat, nuhat ja ihotaudit ovat sairauksia, joista pyydetään tilastoja sekä erikseen että usein myös yhdessä. Kaikista yhteispyynnöistä, joita oli 5 kpl, jaettiin jokainen sairaus omaan ryhmäänsä. Eli pyyntöjen yksittäiset sairaudet otettiin mukaan omina tietoinaan. Yksittäisiä sairauksia oli täten 51 kpl. Kuvasta 9 käy ilmi, että määrällisesti eniten dataa pyydettiin ihotaudeista ja infektioitaudeista (9 kpl molemmista). Myös astmat (8 kpl), syövävät (7 kpl) ja nuhat (5 kpl) ovat olleet suosittuja hakuperusteita. Muita vähemmän pyydettyjä tautiryhmiä ovat asbestisairaudet, rasisairaudet sekä meluvammat. Joitakin yksittäisiä muita sairauksia on myös kysytty, esimerkiksi kivipölykeuhkoja.





Kuva 9. Tietopalvelupyynnöissä esiintyvien sairauksien lukumäärät tautiryhmittäin.

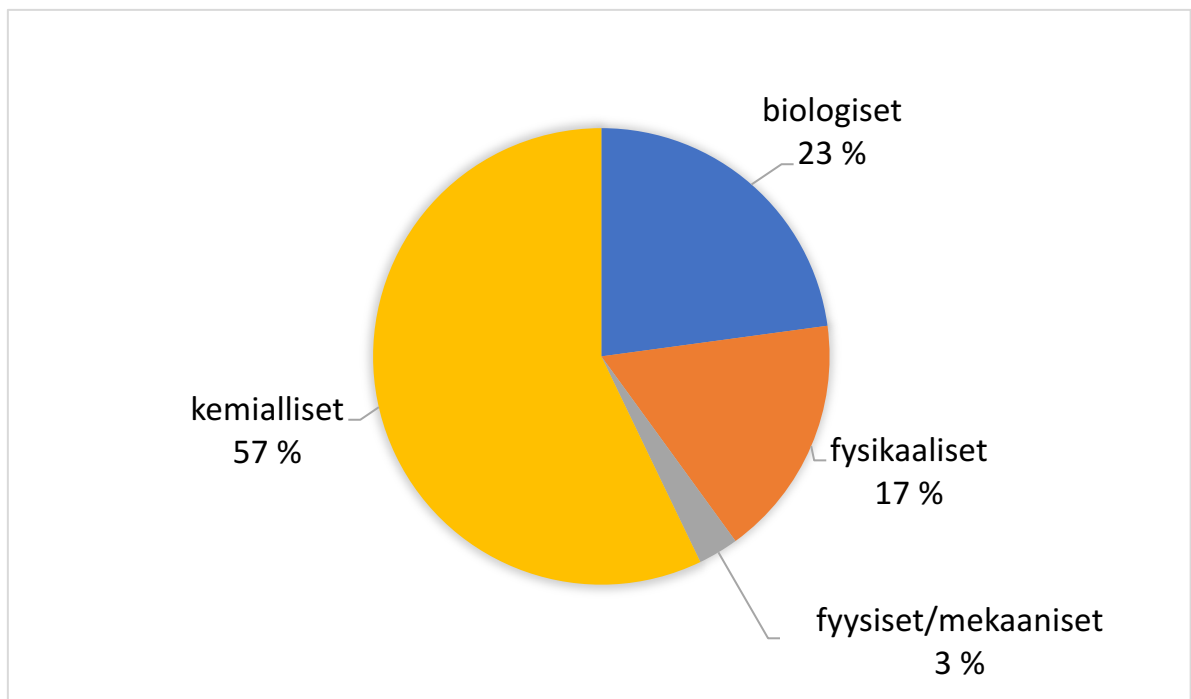
Ristiintaulukoitaessa muuttujat *sairaus* ja pyynnön *hakuehtojen lukumäärä*, havaittiin, että jos pyynnössä oli enemmän kuin yksi hakuehto, oli näistä sairaus 18 pyynnössä mukana toisena hakuehtona, joka on 58 % kaikista kahden hakuehdon pyynnöistä. Sairaus oli mukana viidessä kolmen hakuehdon pyynnössä, ja löytyi myös yksi pyyntö, jossa oli yhteensä neljä hakuehtoa sairauden ollessa niistä yksi.

### 5.2.2 Altiste

Toiseksi suosituimpana hakuehtona rekisterin dataa pyydetessä on altiste. Altisteiden osuus kaikkien pyyntöjen hakuehdoista oli 26 %. Altiste oli mukana 17 pyynnössä (55 %:ssa), kun hakumuuttujia oli kaksi.

Suosituimmat altisteet sijoittuvat kemiallisten altisteiden ryhmään, kuten kuvasta 10 nähdään. Näiden osuus altisteen mukaisista pyynnöistä oli 57 %. Tähän ryhmään kuuluvat muun muassa asbesti, metallit, liimat, maalit, pesuaineet, liuottimet ja kasvipäriset altisteet. Kyseessä on siis laaja ryhmä ja näitä altisteita kohdataan monissa erityyppisissä ammateissa.

Biologisiin altisteisiin liittyviä hakupyynnöitä oli 23 %. Ryhmään kuuluvat muun muassa kosteusvaurioaltisteet sekä monia infektioita aiheuttajia, eli bakteerit, virukset ja alkueläimet. Fysikaalisten altisteiden osuus altisteisiin kohdistuvista pyynnöistä oli 17 %. Nämä pitävät sisällään esimerkiksi melun ja käsitärinän aiheuttamat ammattitautitapaukset. Mekaanisista altisteista (3 %) esimerkiksi toistotyö ja ranteiden poikkeava työasento kuuluvat tähän ryhmään.



Kuva 10. Altisteisiin pohjautuvat tietopalvelupyynnöt altisteryhmittäin.

### 5.2.3 Toimiala

Kolmas merkittävä hakuehto oli toimiala. Toimialojen osuus pyyntöjen hakuehdoista oli 19 %. Toimialapyyntöjen hajanaisuus oli suurta. Taulukkoon 2 on koottu aakkosjärjestyksessä kaikki toimialat, joiden perusteella on tehty tietopalvelupyynnöitä rekisteriin. Kiinnostusta oli moneen eri toimialaan, ja yksittäisistä pyynnöistä vain jokunen oli useammin kuin kerran samasta toimialasta. Useammin kuin kerran pyydettiin tietoja

tapauksista, joiden toimialana on palo- ja pelastustoimi, kasvinviljely ja kotieläintalous ym., elintarvikkeiden ym. valmistus, ravitsemistoiminta, terveydenhuolto, rakentaminen sekä moottoriajoneuvojen huolto ja korjaus ym.

Taulukko 2. Tietopalvelupyynnöissä esiintyneet toimialat (Tilastokeskus TOL2008 mukaan) aakkosittain järjestettynä.

Elintarvikkeiden valmistus
Erikoistunut rakennustoiminta
Hammaslääkäripalvelut
Hammasproteesien, keinohampaiden ym. valmistus
Huoltamatoiminta
Jätteen keruu, käsittely ja loppusijoitus; materiaalien kierrätys
Juomien valmistus
Kampaamo- ja kauneudenhoitopalvelut
Kasvinviljely ja kotieläintalous, riistatalous ja niihin liittyvät palvelut
Kumi- ja muovituotteiden valmistus
Laivojen ja kelluvien rakenteiden rakentaminen
Maa- ja vesirakentaminen
Metsätalous ja puunkorjuu
Moottoriajoneuvojen huolto ja korjaus (pl. moottoripyörät)
Muu rakennus- ja teollisuussiivous
Muu terveyspalvelu
Painaminen ja tallenteiden jäljentäminen
Palo- ja pelastustoimi
Rakennusten ja rakennelmien purku ja rakennuspaikan valmistelutyöt
Ravitsemistoiminta
Sahatavaran sekä puu- ja korkkituotteiden valmistus (pl. huonekalut); olki- ja punonta- tuotteiden valmistus
Siivouspalvelut
Sosiaalihuollon avopalvelut
Sosiaalihuollon laitospalvelut
Talonrakentaminen
Tekstiilien valmistus
Terveyskeskus- ja vastaavat yleislääkäripalvelut
Terveyspalvelut
Vaatteiden valmistus

## 5.2.4 Ammatti

Neljäntenä hakuehtona oli ammatti. Ammatin perusteella haluttiin tietoja 15 %:ssa tapauksista. Aivan kuten toimialat, myös ammatit ovat jakaantuneet laajalle, koska ammatteja ja virallisia ammattikoodeja on paljon. Taulukossa 3 ovat esillä kaikki hakupyynnöissä esiintyneet ammatit. Useammin kuin kerran jonkun tietyn ammatin mukaan pyydettyjä tilastoja on haluttu ammattiryhmistä palomiehet, kampaajat ja kosmetologit, hammashuollon ammatit, esittävät taiteilijat, sekä hoitoalan eri ammateista.

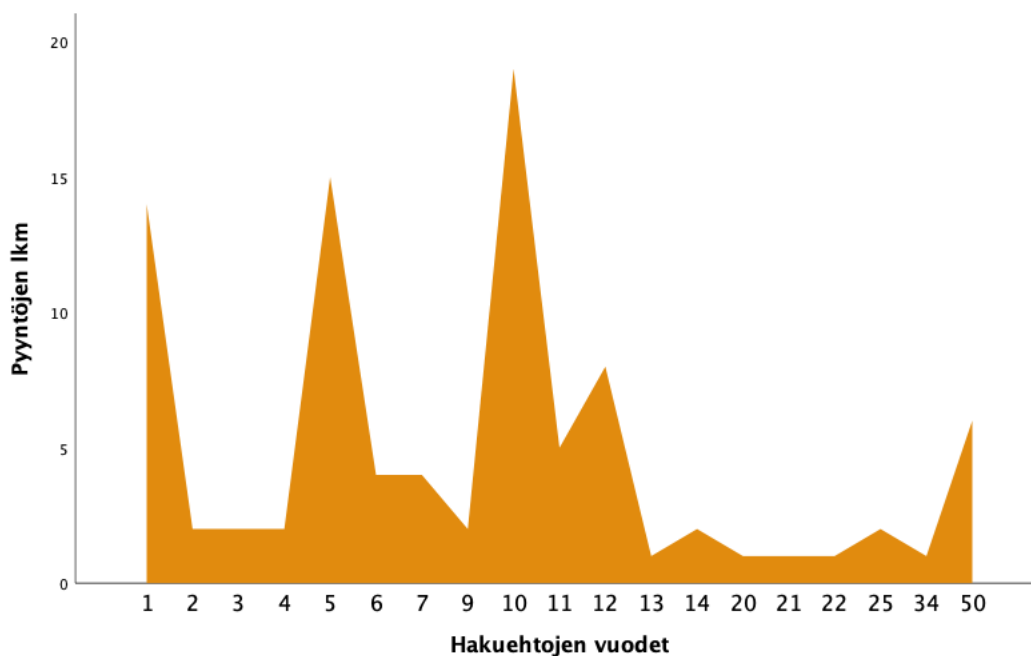
Taulukko 3. Tietopalvelupyynnöissä esiintyneet ammatit (Tilastokeskus AML2010 mukaan) aakkosittain järjestettynä.

Bioanalytikot (terveydenhuolto)
Eläinlääkärit
Eläintenhoitajat ja lemmikkieläinten trimmaajat
Hammas- ja apuvälineteknikot
Hammashoitajat
Hammaslääkärit
Hammasteknikot
Hitsaajat ja kaasuleikkaajat
Huoltamotyöntekijät
Kampaajat ja parturit
Kampaajat, parturit, kosmetologit ym.
Kumituoteteollisuuden prosessityöntekijät
Kuvataiteilijat
Laborantit ym.
Lähihoitajat
Moottoriajoneuvojen asentajat ja korjaajat
Muusikot, laulajat ja säveltäjät
Muut sähköasentajat
Näyttelijät
Ohutlevysepäät
Paksulevysepäät ja rautarakennetyöntekijät
Palomiehet
Sairaanhoidajat ym.
Suuhygienistit
Taiteilijat

## 5.2.5 Pyyntöjen vuosijakauma

Rekisteristä pyydetään tietoa monilla eri hakuehdoilla eli muuttujilla, kuten aiemmissa luvuissa on kerrottu. Yksi tärkeä hakuehto on vuosirajaus rekisterin sisältämiin tietoihin. Tämä ei kuitenkaan ole varsinainen muuttuja, eli tapauksen sisältämä tieto. Tapauksella on yhtenä muuttujana aina rekisteröintivuosi, ja vahvistetulla ammattitaidilla on sen lisäksi vahvistamispäivämäärä. Näitä tietoja käyttäen voidaan rekisteristä etsiä tietoa haluttujen vuosien ajalta.

Kuvasta 11 nähdään, että suosituin aika, jonka sisällä haluttiin dataa, oli 10 vuoden tiedot. Seuraavaksi eniten pyydettiin tietoa 5 vuoden ajalta, tai tuoreimman yhden vuoden ajalta. Jos tietoja haluttiin jonkin muuttujan avulla hakea koko rekisterin olemassaolon ajalta, oli silloin kyseessä haku 50 vuoden tietoihin. Vuosista erottuvat myös luvut 6, 7, 11 ja 12.



Kuva 11. Hakuehtoihin sisältyneet vuosirajaukset ja niiden lukumäärät.

## **5.3 Pyyntöjen sisältö käyttäjäryhmittäin**

Luvussa esitellään tulokset käyttäjäryhmittäin suhteessa tietopalvelupyyntöjen hakuehtoihin. Koska kaksi käyttäjäryhmää eli opiskelijat ja erikoistuvat lääkärit tekivät yhteensä määrällisesti paljon vähemmän pyyntöjä kuin työssäkäyvien ryhmä, tuloksia painotetaan enemmän suhteellisina osuuksina, kuin lukumäärinä.

### **5.3.1 Opiskelijoiden pyynnöt**

Opiskelijoiden ryhmään kuuluvien pyyntöjen määrä oli koko tarkasteluajalta noin 10 % kaikkien pyyntöjen määrästä. Opiskelijoiden pyyntöjen hakuehtona oli useimmiten jokin ammatti (5 kpl), joka oli 25 % kaikkien käyttäjäryhmien pyynnöistä, joissa ammatti oli hakuehtona. Toiseksi eniten opiskelijoiden pyynnöissä haluttiin tietää toimialan mukaista tietoa (4 kpl), joka oli 15 % kaikista toimialan sisältävistä pyynnöistä. Altisteet (2 kpl) ja sairaus (1 kpl) eivät olleet opiskelijoiden pyynnöissä yhtä suosittuja, eli niitä oli mukana vain muutamassa prosentissa suhteutettuna kaikkien pyyntöjen määriin. Jokin muu hakuehto oli mukana yhdessä tapauksessa. Opiskelijoiden pyyntöjen ammatti- ja toimialakiinnostukset kohdistuivat useimmiten hammasteknikoihin, kauneudenhoitoalaan sekä palo- ja pelastustoimeen.

### **5.3.2 Erikoistuvien lääkäreiden pyynnöt**

Erikoistuvien lääkäreiden pyyntöjä oli kaikkien pyyntöjen määrästä 9 %. Heidän pyynnöissään hakuehtoina olivat useimmiten altiste (5 kpl), joka oli kaikista altisteen sisältävistä pyynnöistä 14 %. Yhtä monta kertaa pyynnöissä oli hakuehtona sairaus (5 kpl), joka puolestaan oli 12 % kaikista sairauden sisältävistä pyynnöistä. Jokin muu ehto oli mukana kolmessa pyynnössä. Suosituimpia sairauksia erikoistuvien lääkäreiden pyynnöissä olivat

rasitussairaudet (3 kpl) ja altisteista puolestaan kemiallisten altisteiden ryhmään kuuluvat altisteet (3 kpl).

### **5.3.3 Työssäkävien pyynnöt**

Suurimpana ryhmänä, johon kaikista pyynnöistä kuului 82 %, oli työssäkävien ryhmä. Sekä määrällisesti (36 kpl) että suhteellisesti (86 % kaikista sairauden sisältävistä pyynnöistä) eniten tästä ryhmästä pyydettiin tietoa jonkin sairauden perusteella. Seuraavaksi suosituin hakuperuste oli altiste (28 kpl, 80 % kaikista altisteen sisältävistä pyynnöistä), seuraavaksi toimiala (20 kpl, 77 % kaikista toimialan sisältävistä pyynnöistä) ja ammatti (14 kpl, 70 % kaikista ammatin sisältävistä pyynnöistä). Jokin muu hakuehto oli mukana tämän ryhmän pyynnöissä kahdeksan kertaa, joka oli 67 % kaikista pyynnöistä, joissa oli ”muu hakuehto” mukana.

## 6 JOHTOPÄÄTÖKSET JA ARVIOINTI

Tämän pro gradu -tutkielman tarkoituksena oli selvittää Työperäisten sairauksien rekisterin käyttöä. Rekisterin vastuuhenkilöille neljän vuoden aikana tulleet tietopalvelupyynnöt koottiin aineistoksi, joka työstettiin kvantitatiiviseksi dataksi ja analysoitiin tilastollisin menetelmin. Näiden tutkimustulosten avulla haluttiin saada tietoa, jonka avulla voitaisiin ymmärtää rekisterin käyttöä paremmin, ja sen myötä pystyttäisiin kehittämään tulevia julkaisuja.

Työperäisten sairauksien rekisterin varsinaisesta käytöstä ei ole aikaisemmin tehty analyysiä, joten tämä tutkimus antoi täysin uutta tietoa siitä, miten ja kuka tätä lakisääteisesti ylläpidettävää rekisteriä käyttää. Gissler & Haukka (2004) ovat havainneet aiemmissa rekisteritutkimuksissa vastoinkäymisiä muun muassa siinä, että terveydenhuollon rekistereistä ei ylipäättäen tiedetä tarpeeksi, eikä rekisteritutkimukseen tarvittavia metodologisia taitoja ole riittävästi tutkijoilla tai opiskelijoilla. Tässä pro gradussa ei voitu kysyä tietopalvelupyytäjien kokemuksia tiedon löydettävyydestä tai rekisterin tunnettavuudesta. Joskus tietojen pyytäjät ovat kuitenkin kertoneet, että heidän tarvitsemaansa tietoa on ollut vaikea löytää, eikä moni ollut kuullut TPSR:stä ennen ammattitauteihin liittyvän tiedon etsimistä.

Ensimmäinen tutkimuskysymys selvitti yleisellä tasolla tiedon pyytäjien taustaa. Eniten tietopalvelupyyntöjä teki ryhmä, johon kuuluivat työssäkäyvät henkilöt. Vähemmän pyyntöjä tehtiin opiskelijoiden ja erikoistuvien lääkäreiden taholta. Pughin (1992) ja Fordin (2015) mukaan arkistojen ja asiakirjojen käyttäjien ja kontekstien luokittelujen mukaisesti TPSR:stä tietoja pyytävät henkilöt kuuluvat ammatillisiin käyttäjiin. Heidän tietokäyttäytymisensä kontekstina ovat työ ja opiskelu Fordin (2015) luokittelun mukaan.



Tämä oli selkeä tulos, koska TPSR:n ollessa tutkimusrekisteri, käytetään sen tietoja ammattimaiseen toimintaan, ei harrastukseen tai niin sanottuun vapaa-ajan käyttöön.

Toinen tutkimuskysymys koski tietopalvelupyynnöiden sisältöä ja määrää, eli millaista dataa, ja kuinka paljon mitäänkin tilastoja rekisteristä pyydetään. Pyyntöissä olleet hakuehdot eli muuttujat ryhmiteltiin rekisteristä löytyvien pääasiallisten ja tärkeimpien muuttujien mukaan: sairaus (tauti/tautiryhmä), altiste, toimiala, ammatti, muu hakuehto, vuodet sekä joidenkin näiden kombinaatiot. Tutkimuksessa selvisi, että useimmiten pyyntöjen sisältötoiveisiin kuului jokin sairaus ja toiseksi yleisin oli altisteen mukaan pyydetty data.

Tietopalvelupyynnöissä on usein rajaavana tekijänä jokin aika. Useimmiten tietoa on haluttu 10 tai 5 vuoden ajalta, sekä tuoreimman tilastoidun vuoden eli yhden vuoden ajalta. Vuosirajauksista nousivat esiin myös 6, 7 ja 11, ja ne selittyvät sillä, että pyyntövuosina on haluttu saada tietoja johonkin tiettyyn tasa- tai viisivuotisvuosilukuun saakka taaksepäin. Esimerkiksi jos pyyntö on tehty vuonna 2017, on siitä 12 vuotta vuoteen 2005. Vuosijulkaisuissa esitetään tilastot siis tuoreimman yhden vuoden ajalta, mutta julkaisussa on myös kuvaajia, joissa esitellään tietoja viimeisen 10 vuoden ajalta. Tämä vaikuttaa tulostenkin perusteella edelleen järkevältä rajaukselta. Työelämä-tietopalvelussa puolestaan on esitetty ammattitaudit viiden vuoden ajalta, ja ammattitautitilasto-sivun uudistuksen jälkeen tuota viisivuotisaikaa on mahdollista vaihtaa kolmen eri ajanjakson sisällä (2011-2015, 2012-2016 tai 2013-2017). Viiden vuoden aikajakso oli myös tämän työn tulosten mukaan kiinnostava.

Rekisterin vuosijulkaisu on keskittynyt suureksi osaksi eri sairauksien esittelyyn niin tautiryhmittäin kuin yksittäisinä tarkempina diagnooseina. (Koskela ym. 2021). Tämän pro gradu -tutkimuksen tuloksetkin puoltavat sitä, että sairaudet kiinnostavat eniten, ja niiden

esiintuonti kannattaa jatkossakin. Toki myös altisteryhmiä ja eri sairauksia ammattittain ja toimialoittain esiteltynä on järkevää nostaa esiin vuosittain myös tulevissa julkaisuissa.

Vaikka ammattitauteja (ja niiden epäilyjä) havaitaan vuosittain edelleen paljon meluvam-  
mojen ja asbestiin liittyvien sairauksien osalta (Koskela ym. 2021, 3-4), on huomioitava,  
että niitä todetaan useimmiten vanhemmilla ikäryhmillä, eli 50-vuotiaat ja sitä vanhem-  
mat, ja myös eläkeikäisillä. Meluvamma ja syöpä ovat pitkän latenssin sairauksia, jotka  
tulevat esiin usein vasta vuosia altistumisen jälkeen. Ajankohtaista tietoa nykyhetken am-  
mattitaukeista nimenomaan työikäisten väestössä halutaan selkeästi enemmän, vaikka  
myös syöivistä on pyydetty tilastoja jonkin verran. Meluvamma on vuosijulkaisussa laa-  
jasti esitelty sairaus, koska se on jo pitkään ollut ammattitautien yleisin aiheuttaja. Melu-  
vammoja on kuitenkin pyydetty tämän tutkimuksen mukaan vain kaksi kertaa, joten voi-  
daan olettaa, että vuosijulkaisussa esitettävät meluvammatiedot vaikuttavat olevan riittä-  
vän monipuolisesti esitetty.

Vuosijulkaisussa esitellään kattavasti muun muassa ihotaudit sekä astmat ja nuhat, joita  
on työikäisillä paljon (Koskela ym. 2021, 5-6). Tässä tutkimuksessa huomattiin, että  
vaikka vuosijulkaisusta löytyy melko kattavasti tietoa näistä sairauksista, on niistä silti  
pyydetty paljon erillisiäkin tilastoja. Astman, nuhan tai ihotaudin tai näiden yhdistelmän  
perusteella haluttuja tilastoja oli lähes puolet kaikista sairauden perusteella tehdyistä  
pyynnöistä. Koska nämä ammattitaudit voivat aiheutua monista eri altisteista, vuosijul-  
kaisussa ei voida esittää kaikkia näkökulmia.

Tuloksista nousi esiin ryhmä infektioaudit, josta on pyydetty usein tilastoja. Tilastoja on  
haettu yksittäisten diagnoosikoodien ja/tai altisteiden avulla. Kyseinen tautiryhmä ei ole  
ollut aiemmin vuosijulkaisussa esillä. Infektioaudit sisältävät sairauksia, jotka voivat  
kuulua eri tautiryhmiin, kuten esimerkiksi ihotauteihin. Sekä rekisterin kehittämisen että

vuosijulkaisuissa esitettävien tietojen kannalta infektioitaudeille on syytä luoda oma muuttuja, jonka avulla kyseiset tapaukset saadaan jatkossa nostettua esille.

Sekä toimialojen että ammattien mukaan tehtyjen tietopalvelupyynnöiden hakuehdot liittyivät esimerkiksi palo- ja pelastustoimeen sekä kauneudenhoitoalaan. Näillä aloilla onkin havaittu enenevässä määrin altistavia tekijöitä, ja niiden riskejä on haluttu nostaa esiin, jotta sairauksia voidaan ehkäistä ennen niiden syntymistä (mm. Laiho 2020, 7-8). Myös maatalouden, rakentamisen sekä terveydenhuollon ammateista ja toimialoista on oltu kiinnostuneita. Nämä kaikki nousevat esiin myös vuosijulkaisun tiedoissa, koska niissä havaitaan vuosittain paljon ammattitauteja, mutta vuosijulkaisussa ne tulevat esille aina jonkin sairauden kautta. Vuosijulkaisun rakennetta on yksinkertaistettu viimeisen viiden vuoden aikana, mutta tavoitteena on jatkossa tehdä viiden vuoden välein laajempi katsaus. Tähän katsaukseen olisi hyvä tuoda riittävästi myös toimiala- ja ammattikohtaista tietoa.

Etenkin opiskelijoiden tekemistä pyynnöistä huomattiin, että koska opiskelijat ovat opiskelemassa johonkin tiettyyn ammattiin, he haluavat tietoa nimenomaan kyseiseen ammattiin liittyvistä asioista. Tästä syystä opiskelijoiden pyynnöissä haluttiin eniten saada selville ammateittain esiin tulevia ilmiöitä. Erikoistuvat lääkärit puolestaan tekevät tutkimusta sairauden, eivät juurikaan ammatin kautta. Heitä kiinnostivat eniten erilaiset sairaudet sekä altisteet.

Tietopalvelupyynnöiden tekijäryhmä työssäkäyvät oli määrällisesti suurin. Ja koska ryhmässä oli suhteellisesti eniten pyyntöjä, ryhmän tulokset suosituimmista hakuehdoista myötäilevät pitkälti koko tutkimuksen tuloksia, eli järjestys suosituimman mukaan oli sairaus, altiste, toimiala ja viimeisenä ammatti.

Tuloksissa on huomioitavaa, että tietopalvelupyyntöjen tulo keskeytyi lähes kokonaan samaan aikaan koronapandemian alkamisen kanssa. Tästä voidaan tehdä johtopäätös, että korona vaikutti myös Työperäisten sairauksien rekisterin tietojen käyttämiseen. Koronaviruksesta johtuva sairaus voidaan todeta ammattitaudiksi, mikäli se liitetään työssä saaduksi (Tapaturmavakuutuskeskus 2021c), joten on mielenkiintoista nähdä, kasvaako kiinnostus infektioautien ja biologisten altisteiden tilastointia kohtaan jatkossa.

Tutkimuksen yhdeksi rajoitukseksi huomattiin jo alkuvaiheessa aiheen aiempaan tutkimukseen perehtyessä, että vastaavanlaisia tutkimuksia ei ole tehty. Rekisteritutkimuksia on tehty, mutta ne ovat keskittyneet rekistereiden sisältöjen ja muuttujien arviointiin, sekä vertailuun eri rekisterien kesken, mutta ei varsinaisesti rekisterin käyttöön liittyviin asioihin. Tutkimuksia löytyy myös rekistereiden sisällöllisistä asioista, mikä puolestaan ei ollut tämän työn mielenkiinnon kohde. Tässä pro graduussa huomioitiin kuitenkin ne tutkimukset, joiden aiheet liittyivät vastaavanlaisiin työperäisten sairauksien rekistereihin.

Varsinaisia tutkimuksia rekisterien käyttäjistä ei löytynyt, joten työssä on keskitytty arkistojen käyttäjiin liittyviin tutkimuksiin. Teoreettisen viitekehysten pohdinnassa selvitettiin, että Työperäisten sairauksien rekisteri täyttää joiltakin osin asiakirjallisen aineiston määritelmän, joten arkistojen ja asiakirjojen käyttäjiin suuntautuvat tutkimukset sitovat aiheen myös TPSR:n käyttöön. Pro graduun alun perin suunnitellun käyttäjille suunnatun kyselytutkimuksen poisjäänti aiheutti sen, että arkistojen käyttäjiin viittaavaan kirjallisuuteen tehdyn katsauksen näkökulma ja yhteys liittyen tutkimuskysymykseen rekisterin käyttäjistä ja heidän tietokäyttämistään jäi valitettavan niukaksi. Tiedot käyttäjien kokemuksista esimerkiksi vuosijulkaisun sisällön ja Työelämä-tieto-palvelun käytettävyyden osalta jäivät siis tässä työssä toteutumatta.

Tästä pro gradu -tutkielmasta pois jätetty osio eli kyselytutkimus rekisterin käyttäjille on jatkotutkimuksena erittäin mahdollinen ja toivottu. Kysely voitaisiin suunnata sekä tietopalvelupyyntöjen tekijöille että suuremmalle joukolle rekisteritietojen käyttäjiä, eli yhteiskunnallisesti vaikuttaville tahoille. Näitä ovat muassa sosiaali- ja terveystieteiden ministeriö, Tapaturmavakuutuskeskus, Maatalousyrittäjien eläkelaitos, vakuutusyhtiöt, ammattiyhdistykset sekä muut yhteiskunnalliset toimijat. Jatkossa myös tämän tutkimuksen mukainen käyttäjä- ja käytön analyysi olisi hyvä tehdä säännöllisin väliajoin. Siten voitaisiin tuottaa tietoa, miten ammattitautitilastojen käyttö muuttuu ajan myötä.

Tämä tutkimus oli ensimmäinen rekisterin käytöstä koskaan tehty analyysi ja tavoite rekisterin käytön paremmasta ymmärryksestä toteutui, koska tutkimuksella saatiin tärkeää uutta tietoa rekisterin käytöstä. Työelämä-tietopalvelun ammattitautiosio vastaa osittain tilastotarpeeseen, mutta sieltä ei ole mahdollista saada yhtä spesifejä tietoja kuin tekeillä rekisteritiimille tietopalvelupyntöjä. Tuloksista päätellen vuosijulkaisua yksityiskohtaisemman ammattitautitiedon tarvitsijoita on melko paljon ja heitä on monilla eri osa-alueilla niin työelämässä kuin opintojen parissa. On siis tärkeää jatkossakin pystyä vastaamaan tietopalvelupyntöihin asiantuntevan rekisteritiimin avulla.

## LÄHTEET

Alastalo, M. & Borg, S. (2010). Numerolukutaito. Kvantitatiivisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoaarkisto. Saatavilla: <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvanti/numerolukutaito/analyysi/>

Borglund, E. (2005). Operational use of electronic records in police work. Information Research 10(4), paper 236. Saatavilla: <http://informationr.net/ir/10-4/paper236.html>

Briney, K. (2015). Data management for researchers: organize, maintain and share your data for research success.

Brookhart, M. A., Stürmer, T. & Glynn, R. J. (2010). Confounding Control in Healthcare Database Research: Challenges and Potential Approaches. Medical Care, 06/2010, Vuosikerta / Sarjan osa 48, Numero 6 Suppl 1.

Case, D. & Given, L. (2016). Looking for Information: a survey of research on information seeking, needs and behaviour. 4th edn, Emerald.

Davoodi, S., Haghghi, K. S. & Kalhori, S. R. N. (2017). Occupational Disease Registries—Characteristics and Experiences. ACTA INFORM MED. 2017 JUN; 25(2): 136-140. doi: 10.5455/aim.2017.25.136-140

Edwards, P.N., Jackson, S.J., Bowker, G.C. & Knobel, C.P. (2007). Understanding Infrastructure: Dynamics, Tensions, and Design. Final report of the workshop History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures, NSF, Office of Cyberinfrastructure. Saatavilla: <http://hdl.handle.net/2027.42/49353>

EUR-Lex. Access to European Union law. (2021). Regulation (EC) No 1338/2008. Saatavilla: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32008R1338>

Findata. (2020). Saatavilla: <https://www.findata.fi/>

Ford, N. (2015). Introduction to Information Behaviour.

Gissler, M. & Haukka, J. (2004). Finnish health and social welfare registers in epidemiological research. *Norsk Epidemiologi* 2004; 14 (1): 113-120.

Haasio, A., Harviainen, T. J. & Savolainen, R. (2019). Johdatus tiedonhankintatutkimukseen. Helsinki: Avain.

Heikkilä, T. (2014). Tilastollinen tutkimus. Helsinki: Edita Publishing Oy.

Henttonen, P. (2015). Johdatus asiakirjahallinnan tutkimukseen. Helsinki: Avain.

Hirsjärvi, S., Remes, P. & Sajavaara, P. (2015). Tutki ja kirjoita. 20. painos. Helsinki: Tammi.

Hughes, L. M. (2012). Evaluating and Measuring the Value, Use and Impact of Digital Collections.

Jackson, S. J., Edwards, P. N., Bowker, G. C. & Knobel, C. P. (2007) Understanding Infrastructure: history, heuristics, and cyberinfrastructure policy, *First Monday*,12(6). Saatavilla: <https://www.firstmonday.org/ojs/index.php/fm/article/view/1904/1786>

Johnson J. D. (2014). Health-related information seeking: Is it worth it? *Information Processing and Management* 50 (2014) 708–717. Saatavilla: <http://dx.doi.org/10.1016/j.ipm.2014.06.001>

Karasti, H., Janasik, N., & Talja, S. (2011). Tieteen infrastruktuurit – näkymättömiä tukirakenteita, kansallisia kilpailuvaltteja vai uuden tutkimusvision mahdollistajia? Tieteessä Tapahtuu, 29(3). Saatavilla: <https://journal.fi/tt/article/view/4097>

Koskela, K., Aalto-Korte, K., Pesonen, M., Suojalehto, H., Lindström, I., Airaksinen, L., Suuronen, K., Lehtimäki, J., Toivio, P. & Helaskoski, E. (2019). Riskiammatit, -alat ja -altisteet työperäisten sairauksien rekisterin perusteella: 10 vuoden aineiston analyysi. Helsinki: Työterveyslaitos. Saatavilla: <http://urn.fi/URN:ISBN:9789522618580>

Koskela, K., Lehtimäki, J., Aalto-Korte, K., Pesonen, M., Lindström, I., Suojalehto, H., Airaksinen, L., Suuronen, K. & Helaskoski, E. (2020). Ammattitaudit ja ammattitautiepäilyt 2016: Työperäisten sairauksien rekisteriin kirjatut uudet tapaukset. Helsinki: Työterveyslaitos. Saatavilla: <http://urn.fi/URN:ISBN:9789522619273>

Koskela, K., Lehtimäki, J., Aalto-Korte, K., Pesonen, M., Lindström, I., Suojalehto, H., Airaksinen, L., Suuronen, K. & Helaskoski, E. (2021). Ammattitaudit ja ammattitautiepäilyt 2017: Työperäisten sairauksien rekisteriin kirjatut uudet tapaukset. Helsinki: Työterveyslaitos. Saatavilla: <http://urn.fi/URN:ISBN:9789522619280>

Laiho, S. (2020). Hiusalan kemialliset altisteet ja työperäisen oireilun ehkäisy. Laurea-ammattikorkeakoulu. Saatavilla: <http://urn.fi/URN:NBN:fi:amk-2020060216358>

Lybeck, J. et al. (2006). Arkistot yhteiskunnan toimiva muisti. Helsinki: Arkistolaitos.

Mellin, I. (2006). Tilastolliset menetelmät: Johdanto. Saatavilla: <https://math.aalto.fi/opetus/sovtoda/oppikirja/Johdanto.pdf>

Mäkivaara, T. (2017). Tietopalvelupyyntöjen hakutermien analyysi yksityisen organisaation kontekstissa. Tampereen yliopisto. Viestintätieteiden tiedekunta,



informaatiotutkimus ja interaktiivinen media. Pro gradu -tutkielma. Saatavilla: <http://urn.fi/URN:NBN:fi:uta-201801051023>

Pajari, J. (2019). Tutkimusaineistojen metatiedot. Metatietojen laatu data- ja metatietoarkistoissa. Tampereen yliopisto. Viestintätieteiden tiedekunta, informaatiotutkimus ja interaktiivinen media. Pro gradu -tutkielma. Saatavilla: <http://urn.fi/URN:NBN:fi:tuni-202003172724>

Pryor, G. (2012). Managing Research Data. London: Facet Publishing. eBook ISBNs: 9781856047562. 9781856048910. EBSCOhost AN 558459.

Pugh, M. J. (1992). Providing Reference Services For Archives And Manuscripts. Chicago: Society of American Archivists.

Ruokatieto. (2021). Tilastoja – Tietohaarukka. Saatavilla: <https://www.ruokatieto.fi/ruokafakta/tietohaarukka-kokoa-luvut-pelloilta-poytaan>

Savolainen, R. (2012). Conceptualizing information need in context. Information Research, 17(4) paper 534. Saatavilla: <http://InformationR.net/ir/17-4/paper534.html>

Schellenberg, T. R. (1956). (Reprint 1998). Modern Archives: Principles and Techniques. The Society of American Archivists. Kansas State Historical Society. ISBN 0-931828-49-X.

Sundqvist, A. (2007). The use of records – a literature review. Archives & Social Studies, 1(1), 623-653.

Sundqvist, A. (2009). Search processes, user behaviour and archival representational systems. Mid Sweden University.

Talja, S., Karasti, H. & Janasik, N. (2010). Tutkimusinfrastruktuuripolitiikan haasteet Suomessa: Tieteen infrastruktuurit -asiantuntijapaneelin antia. Informaatiotutkimus 29 (4), 2010.

Tapaturmavakuutuskeskus. (2021a). Mikä on ammattitauti? Saatavilla: <https://www.tvk.fi/korvaaminen/ammattitauti/>

Tapaturmavakuutuskeskus. Tapaturma-asiain korvauslautakunta. (2021b). Saatavilla: <https://www.tvk.fi/tapaturmavakuutuskeskus/tapaturma-asiain-korvauslautakunta/>

Tapaturmavakuutuskeskus. Koronaviruksen aiheuttaman sairauden korvaaminen ammattitautina. (2021c). Saatavilla: <https://www.tvk.fi/uutiset-ja-blogit/uutiset/2020/koronaviruksen-aiheuttaman-sairauden-korvaaminen-ammattitautina/>

Terveystieteiden tutkimuskeskus ja hyvinvoinnin laitos. (2011). Tautiluokitus ICD-10. Saatavilla: <http://urn.fi/URN:NBN:fi-fe201205085423>

Tieteen termipankki. (2020a). Avoin tiede: rekisteri. Saatavilla: [https://tieteentermipankki.fi/wiki/Avoin\\_tiede:rekisteri](https://tieteentermipankki.fi/wiki/Avoin_tiede:rekisteri)

Tieteen termipankki. (2020b). Avoin tiede: tutkimusaineisto. Saatavilla: [https://tieteentermipankki.fi/wiki/Avoin\\_tiede:tutkimusaineisto](https://tieteentermipankki.fi/wiki/Avoin_tiede:tutkimusaineisto)

Tietosuojavaltuutetun toimisto. (2020a). Euroopan tietosuojaneuvoston ohjeet. Saatavilla: <https://tietosuoja.fi/euroopan-tietosuojaneuvoston-ohjeet>

Tietosuojavaltuutetun toimisto. (2020b). Henkilötietojen käsittely. Saatavilla: <https://tietosuoja.fi/henkilotietojen-kasittely>

Tilastokeskus. (2021). Luokitukset. Saatavilla: <https://www.stat.fi/luokitukset/>

Tough, A. (2006). Record Keeping in a Hybrid Environment: Managing the Creation, Use, Preservation and Disposal of Unpublished Information Objects in Context.

Tutkimuseettinen neuvottelukunta. (2019). Hyvä tieteellinen käytäntö. Saatavilla: <https://www.tenk.fi/fi/hyva-tieteellinen-kaytanto>

Työelämätieto. (2020). Saatavilla: <https://www.tyoelamatieto.fi/#/fi/>

Työterveyslaitos (2020). Työperäisten sairauksien rekisteri. Saatavilla: <https://www.ttl.fi/rekisterit/tyoperaisten-sairauksien-rekisteri/>

Valli, R. & Aaltola, J. (toim.). (2015). Ikkunoita tutkimusmetodeihin 1. Jyväskylä: PS-kustannus.

Vastuullinen tiede. (2020). Hyvä tieteellinen käytäntö. Saatavilla: <https://vastuullinentiede.fi/fi/tutkimustyo/hyva-tieteellinen-kaytanto>

Vilka, H. (2007). Tutki ja mittaa. Määrällisen tutkimuksen perusteet. Helsinki: Kustannusosakeyhtiö Tammi.