

Mikke Bergström

**PUHEENTUNNISTUSTEKNOLOGIAN KÄYTTÖ
TERVEYDENHUOLLON POTILASKERTOMUKSISSA**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2020

TIIVISTELMÄ

Bergström, Mikke

Puheentunnistusteknologian käyttö terveydenhuollon potilaskertomuksissa

Jyväskylä: Jyväskylän yliopisto, 2020, 30 s.

Tietojärjestelmätiede, kandidaatin tutkielma

Ohjaaja: Clements, Kati

Terveydenhuollon käyttömenojen jatkuva nousu pakottaa palveluntarjoajat kehittämään alati kustannustehokkaampia ja laadukkaampia ratkaisuja. Puheentunnistusteknologian nopea kehittyminen mahdollistaa sen käytön useissa terveydenhuollon järjestelmissä ja palveluissa. Tässä kirjallisuuskatsauksena toteutetussa kandidaatin tutkielmassa käsitellään puheentunnistusteknologian käyttöä terveydenhuollon potilaskertomusten saneluissa. Aihe on tärkeä, sillä puheentunnistusteknologian uusilla läpimurroilla voidaan saavuttaa entistä kustannustehokkaampi, laadukkaampi sekä turvallisempi tapa sanella potilaskertomuksia. Tutkielmassa perehdytään asialle olennaisiin käsitteisiin ja teknologioihin kuten terveydenhuoltoon, sen järjestelmiin, potilassaneluihin ja puheentunnistukseen. Tutkimuskysymyksinä ovat *"Miten puheentunnistusteknologiaa hyödynnetään terveydenhuollon potilaskertomuksissa?"* ja *"Mitä hyötyjä puheentunnistusteknologian käyttö voi saavuttaa terveydenhuollon potilaskertomuksissa?"* Tutkielman lähteet on haettu Google Scholar ja JYK-DOK hakupalveluista. Lähteiden osalta painotettiin mahdollisimman laadukkaan ja ajantasaisen tiedon hankintaan. Tutkimuksessa selvisi, että puheentunnistusta on käytetty terveydenhuollon potilaskertomuksissa jo useita vuosikymmeniä. Kuitenkin vasta 2010-luvulla sen potentiaalia on saatu hyödynnettyä entistä enemmän. Puheentunnistuksen käytön hyötyjä ovat sen tuomat säästöt kustannuksissa, tehokkuuden lisääntyminen ja potilaskertomusten eheyden parantuminen. Puheentunnistus ei ole kuitenkaan vielä täydellinen ja jatkotutkimusaiheina olisi hyvä tutkia sen todellisia kustannushyötyjä, käytön hyväksymistä käyttäjien keskuudessa sekä tunnistustarkuutta huippumoderneilla teknologioilla.

Asiasanat: automaattinen puheentunnistus, terveydenhuoltojärjestelmä, potilassanelu, potilaskertomus, tehokkuus

ABSTRACT

Bergström, Mikke

Application of speech recognition in healthcare medical records

Jyväskylä: University of Jyväskylä, 2020, 30 p.

Information Systems, Bachelor's Thesis

Supervisor: Clements, Kati

The constant rise of healthcare operating costs is forcing service providers to develop more cost-effective and high-quality solutions. The rapid development of automatic speech recognition technology also known as ASR enables its use in many healthcare systems and services. This Bachelor's thesis, which is a literature review, addresses the use of speech recognition technology in dictation of medical records. The topic is important because new breakthroughs in speech recognition technology can provide a more cost-effective, high-quality, and safer way to dictate medical records. The thesis explores relevant concepts and technologies such as electronic healthcare systems, medical transcriptions and automatic speech recognition. The research questions are "How is speech recognition technology utilized in healthcare medical records?" and "What benefits can the use of speech recognition technology have in dictating medical records?" The sources for the thesis are from Google Scholar and JYK-DOK search services. While researching the sources I focused on finding the highest quality and up-to-date information about the subject. The study found that speech recognition has been used in used healthcare medical records for several decades. However, it is not until the 2010s that its potential has been fully exploited. Instead, the benefits of using voice recognition are cost savings, increased efficiency, and improved patient record integrity. However, speech recognition is not perfect and potential research topics would be real cost benefits, acceptance among users and recognition accuracy with state of the art technologies.

Keywords: automatic speech recognition, health care system, dictation, medical record, efficiency.

KUVIOT

KUVIO 1 Tyypillisen puheentunnistusjärjestelmän kaavio (Gruhn ym., 2011).	10
KUVIO 2 Kolmen tilan Markovin piilomalli (Gruhn ym., 2011)	11
KUVIO 3 Potilaskertomuksen saneluprosessi.....	16
KUVIO 4 Potilassaneluiden puheentunnistusjärjestelmä.....	19

TAULUKOT

TAULUKKO 1 Puheentunnistusratkaisujen vertailu.....	21
---	----

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
KUVIOT	4
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	6
2 PUHEENTUNNISTUS	8
2.1 Puheentunnistus yleisesti	8
2.2 Automaattinen puheentunnistus	9
2.2.1 Akustinen malli	11
2.2.2 Kielimalli	11
2.3 Puheentunnistuksen käyttötarkoitukset	11
2.4 Puheentunnistusteknologian uudet ratkaisut	12
3 TERVEYDENHUOLLON POTILASKERTOMUKSET	14
3.1 Terveydenhuoltojärjestelmä.....	14
3.2 Potilaskertomukset ja niiden sanelu	15
3.3 Perinteisen potilassanelun haasteet	16
4 PUHEENTUNNISTUS TERVEYDENHUOLLOSSA.....	18
4.1 Puheentunnistus pohjainen potilassanelu	18
4.2 Puheentunnistusjärjestelmän hyödyt	21
4.3 Puheentunnistusjärjestelmän haittapuolet.....	23
5 YHTEENVETO	25
LÄHTEET	27

1 JOHDANTO

OECD Health Statistics (2019) tilaston mukaan näemme terveydenhuollon käytömenojen olleen jatkuvassa kasvussa ympäri maailmaa jo useiden vuosikymmenten ajan. Tästä johtuen terveydenhuollon palveluntarjoajat ovat tehneet valtavia investointeja terveydenhuollon kustannustehokkuuden ja laadun parantamiseen sen lukuisissa osa-alueissa. Yksi näistä investoinneista on jo useassa maassa käyttöönotettu sähköinen terveystietomusarkisto (Electronic Health Records, EHRs), joka tähtää merkittäviin lyhyen ja pitkän tähtäimen säästöihin. EHR-järjestelmien hyödyt ovat kiistattomat. Näitä ovat mm. vähentyneet hoitovirheet, parantunut tiedonvaihto sairaaloiden välillä sekä pienemmät kustannukset terveydenhuollon kansallisella tasolla (Kruse ym., 2016). Terveystietomusten eli toisin sanoen potilaskertomusten kirjaus EHR-järjestelmiin on kuitenkin usein työläs ja aikaa vievä prosessi. Kasvavat potilasmäärät, hoitohenkilöstön työvoimapula ovatkin johtaneet potilaskertomusten työkulun jatkuvaan tehostamiseen.

Puheentunnistusteknologia on saanut paljon huomiota terveydenhuollon dokumentoinnin työkaluna. Puheentunnistuksella tarkoitetaan teknologiaa, jonka avulla tietokone pystyy purkamaan käyttäjän sanelemaa puhetta tekstiksi (Chavan, Ganesh, Sable, 2013). Puheentunnistusjärjestelmien teknologiset rajoitteet ovat asettaneet kuitenkin haasteita ihmisperheen tarkalle ymmärtämiselle. 2010-luvulla tapahtunut tekoälyn ja neuroverkkojen huima kehitys on avannut puheentunnistukselle ovia, jotka ennen olivat mahdottomia saavuttaa. Tämän seurauksena puheentunnistuksen integrointi EHR-järjestelmiin on kasvanut merkittävästi, joka on parantanut potilaskertomusten dokumentointia, tehostanut toimintaa ja mahdollistanut selviä kustannussäästöjä (Saxena ym., 2018). Koska terveydenhuoltojärjestelmät kattavat lukemattomia osa-alueita, keskitytään tässä tutkielmassa vain potilaskertomusten luomiseen ja vertailuun puheentunnistusteknologiaa käyttäen. Tutkielman tavoitteena on avata puheentunnistusta teknologiana ja luoda selkeä kuva siitä, miten sitä hyödynnetään nykyaikaisten terveydenhuoltojärjestelmien potilasasiakirjojen dokumentoinnissa. Tämän tutkielman johtavina tutkimuskysymyksinä ovat:

- Miten puheentunnistusteknologiaa hyödynnetään terveydenhuollon potilaskertomuksissa?
- Mitä hyötyjä puheentunnistusteknologian käyttö voi saavuttaa terveydenhuollon potilaskertomuksissa?

Tämä tutkielma on tehty kirjallisuuskatsauksena ja sen lähteitä on haettu Google Scholar hakupalvelusta sekä Jyväskylän yliopiston JYKDOK-tutkimustietokannasta. Aineistojen haussa on keskitytty siihen, että tutkimuskysymykseen saadaan vastattua mahdollisimman ajantasaisella tiedolla unohtamatta lähteiden luotettavuutta. Puheentunnistusteknologian nopea kehittyminen loi kuitenkin haasteita relevanttien lähteiden etsimiselle. Koska itse puheentunnistusteknologiaa on kuitenkin käytetty jo pitkään, käytetään sitä käsittelevässä kappaleessa myös hieman vanhempia lähteitä.

Tutkielma on rajattu neljään osaan muodostuen johdannon lisäksi neljästä sisältöluvusta. Ensimmäisessä sisältöluvussa käydään läpi itse puheentunnistusteknologiaa. Kappale pohjustaa puheentunnistusta raapaisemalla nopeasti sen historiaa ja kertoen sen teknologiasta ja kehityksestä nykyaikana. Toisessa sisältöluvussa käydään läpi terveydenhuoltoa ja sen järjestelmiä. Samassa luvussa tutustutaan myös potilaskertomusten tuottamiseen potilassaneluiden muodossa. Kolmannessa luvussa käydään läpi ensimmäisen ja toisen luvun pohjustettuja aiheita, eli siis puheentunnistuksen käyttöä terveydenhuollon potilaskertomusten luomisessa. Kyseisessä luvussa pyritään vertailemaan kyseisen teknologian hyötyjä ja haittoja mahdollisimman objektiivisesti. Neljännessä ja viimeisessä luvussa käydään läpi yhteenvedona tutkielman pohjalta tehtyjä havaintoja ja päätelmiä, sekä ehdotetaan jatkotutkimusaiheita.

2 PUHEENTUNNISTUS

Puheentunnistusteknologialla on pitkä historia niin tieteen kuin populaarikulttuurin osalta. Ihmisten tahto automatisoida jokapäiväisiä toimintoja, johti myös kiinnostukseen opettaa tietokone ymmärtämään ihmisen puhetta (Juang & Rabiner, 2005). Puheentunnistuksella tarkoitetaan teknologiaa, jonka avulla tietokone pystyy tulkitsemaan käyttäjän mikrofonilla sanelemaa puhetta tekstiksi (Chavan ym., 2013). Tämä teksti taas muodostetaan sanoiksi, joita tietokone pystyy hyödyntämään useissa käyttötarkoituksissa. Yleisimpänä näistä valmiin tekstin tuottaminen puheesta (Chavan ym., 2013).

2.1 Puheentunnistus yleisesti

Aikaisimmat puheentunnistusteknologiat loivat pohjaa jo 1950-luvulla, kun Bell Laboratoriorien tutkijat loivat järjestelmän numeroiden tunnistamiseen yksittäisen henkilön puheesta. Teknologia otti useita harppauksia eteenpäin seuraavina vuosikymmeninä, mutta nykyisen kaltaisia Markovin piilomalliin ja neuroverkkoihin pohjautuvia puheentunnistusratkaisuja nähtiin vasta 1990-luvun alussa. (Juang & Rabiner, 2005.)

Puheentunnistusratkaisut voidaan luokitella joko puhujasta riippuviin tai riippumattomiin järjestelmiin. Puhujasta riippuvat järjestelmät vaativat käyttäjiltään yksilöllisen äänimallin luomisen sekä sen harjoittamisen (Gales & Young, 2008). Kyseiset järjestelmät ovat yleensä helpommin kehitettäviä, halvempia ja tunnistukseltaan tarkempia. Vastapainona ne eivät ole läheskään yhtä joustavia kuin puhujasta riippumattomat järjestelmät sillä niiden tunnistustarkkuus vaihtelee merkittävästi eri puhujilla (Khilari & P., 2015). Puhujasta riippumattomat järjestelmät taas toimivat ilman yksilöllistä äänimallia ja kaikilla käyttäjillä (Gales & Young, 2008). Ne tunnistavat suurien ihmisjoukkojen puhettavat ja tarjoavat kaikista eniten joustavuutta. Tämän takia ne ovat hankalampia ja kalliimpia kehittää sekä ne tarjoavat yleisesti huonomman tunnistustarkkuuden (Khilari & P., 2015).

Puheentunnistusjärjestelmät voidaan Khilarin ja V.P:n (2015) mukaan jaotella luokittain sen mukaan mitä lausuntatapoja ne tunnistavat :

- Eristetty sana (Isolated Words): Eristettyjä sanoja tunnistavat järjestelmät vaativat tauon eli tietyn verran hiljaisuutta tunnistettavien sanojen välillä. Järjestelmä tunnistaa vain yksittäisiä sanoja tai lyhyitä lausuntoja.
- Yhdistetty sana (Connectec Word) : Yhdistettyjä sanoja tunnistavat järjestelmät ovat samankaltaisia eristettyihin sanoja tunnistaviin järjestelmiin, mutta ne mahdollistavat pienemmän tauon sanojen väliin.
- Jatkuva puhe (Continious Speech): Jatkuvan puheen järjestelmät mahdollistavat käyttäjän puhumisen lähes luonnollisesti samalla kun kone käsittelee tunnistettua puhetta taustalla. Kyseiset järjestelmät ovat hankalimpia toteuttaa sillä ne käyttävät monimutkaisia menetelmiä määrittämään lausuntojen rajat.
- Spontaani puhe (Spontaneous Speech): Spontaania puhetta tunnistavat järjestelmät tunnistavat puhetta jota voidaan pitää käytännössä täysin luonnollisena. Kyseisten järjestelmien tulisi pystyä käsittelemään puheen luonnollisia vivahteita ja epärytmikästä puhetta.

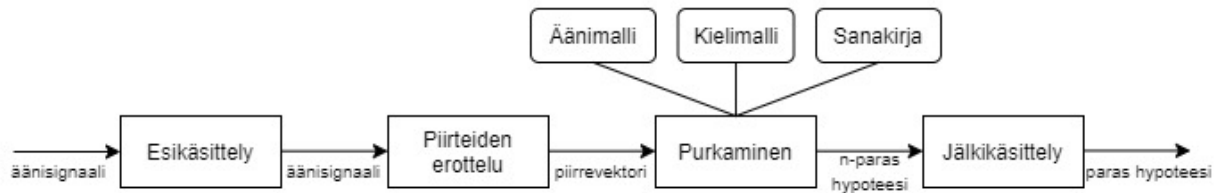
2.2 Automaattinen puheentunnistus

Puheentunnistus joka tunnetaan myös nimellä automaattinen puheentunnistus (automatic speech recognition, ASR), tarkoittaa käytännössä ihmisäänen ymmärrystä tietokoneen avulla (Saini & Kaur, 2013). Ymmärtääksemme puheentunnistusteknologiaa, tulee meidän ensiksi ymmärtää mistä ihmispuhe koostuu. Ihmispuhe koostuu lukuisista yksilöllisistä äänistä eli foneemeista (Gruhn, Minker, Nakamura, 2011). Foneemi on puheen pienin ääniyksikkö, joka erottaa yhden sanan toisesta ja nämä yksiköt vaihtelevat kielittäin (Gruhn ym., 2011). Tämän lisäksi ihmispuhe sisältää hyvin kieli- ja puhujakohtaisia prosodeja eli puheen rytmiä, painotusta, sointia ja voimakkuutta sisältäviä muuttujia (Gruhn ym., 2011). Automaattisten puheentunnistusjärjestelmien tehtävänä on muuttaa ihmispuhe tietojärjestelmän luettavaksi äänisignaalksi ja päätellä siinä tuotetut sanat (Gruhn ym., 2011).

Gruhn ym., (2011) mukaan tyypillinen puheentunnistusjärjestelmä voidaan jaotella seuraaviin toimenpiteisiin kuvion 1 mukaan.

- Äänisignaalin esikäsittely (preprocessing)
- Puheen piirteiden erottelu (feature extraction)
- Puheen purkaminen (decoding) hyödyntäen ääni- ja kielimallia sekä sanavarastoa

- Tulosten jälkikäsitteily ja esittely (post processing)



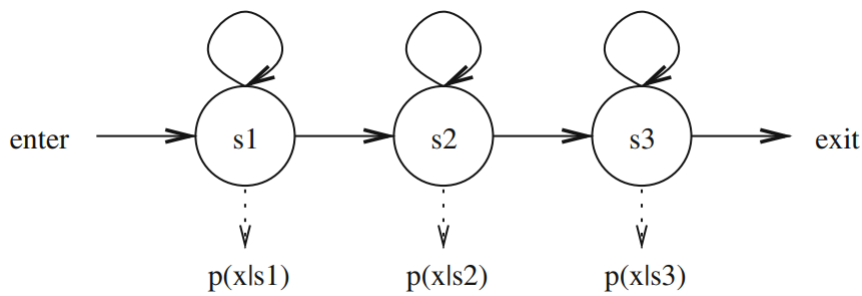
KUVIO 1 Tyypillisen puheentunnistusjärjestelmän kaavio (Gruhn ym., 2011)

Äänisignaalin esikäsitteilyn tarkoituksena on muuntaa ääninäytteet analogisesta muodosta tietokoneelle luettavaan digitaaliseen muotoon (Saini & Kaur, 2013). Äänisignaalista pyritään nyt poistamaan mahdollinen taustamelu, jonka jälkeen se jaetaan n. 25ms aikaikkunoihin puheelle tyypillisten piirteiden erottelua varten (Saini & Kaur, 2013). Piirteiden erottelu pyrkii tarjoamaan kompaktin esityksen akustisen signaalin aaltomuodoista järjestelmälle hyödynnettävässä muodossa (Gales & Young, 2008). Kyseisen vaiheen tarkoituksena on minimoida sanojen välillä syntyvän informaation kato ja täten tarjota hyvä pohja äänisignaalin purkamista ja tulkitsemista varten (Gales & Young, 2008).

Äänisignaalin purkamisen aikana laskelmoidaan mikä sanasarja vastaa todennäköisimmin vastaanotettua akustista signaalia. Gruhn ym., (2011) mukaan puheen purkaminen vaatii kolme tietolähdettä.

- Akustisen mallin sisältäen Markovin Piilomallin (HMM) mallinnetuna jokaiselle yksikölle kuten foneemille tai sanalle (acoustic model)
- Kielimallin (language model)
- Sanakirjan (pron. dictionary)

Lähes kaikki modernit puheentunnistusjärjestelmät hyödyntävät Markovin piilomalliin (HMM) pohjautuvia ratkaisuja puheen purkamista varten (Gales & Young, 2008). Markovin piilomalli on tilastollinen malli, jonka avulla pyritään ennustamaan tuntemattomien (piilotettujen) muuttujien arvo havaittujen muuttujien joukosta (kuviot 2), eli tässä tapauksessa äänitetystä puheesta (Chavan ym., 2013). Koska jokaisella HMM-tilalla voi olla ääretön määrä mahdollisia lopputulemia, hyödynnetään Viterbi-algoritmia laskemaan niistä kaikista todennäköisin lopputulos (Saini & Kaur, 2013). Viterbi-algoritmi antaa yleensä noin 5-10 todennäköisintä hypoteesia käsittelemästään äänisignaalista, jotka se listaa todennäköisyysjärjestykseen (Gruhn ym., 2011).



KUVIO 2 Kolmen tilan Markovin piilomalli (Gruhn ym., 2011)

2.2.1 Akustinen malli

Kuten aikaisemmin mainittiin ihmispuhe koostuu yksittäisistä foneemeista. Akustisella mallilla tarkoitetaan tilastollista mallia puheen foneemeista, jotka muodostavat oikeita sanoja (Gales & Young, 2008). Akustinen malli voi myös sisältää erilaisia mallinnuksia ihmisille tyypillisistä puhetavoista, kuten hengityksestä, maiskuttelusta ja ympäristön äänistä (Gruhn ym., 2011). Puheentunnistuksessa käytettävillä akustisilla malleilla vertaillaan foneemeja vastaanotettuun audiosignaaliin ja lasketaan näin todennäköisyyksiä äänitetystä sanasta yhdessä kielimallin kanssa (Gruhn ym., 2011).

2.2.2 Kielimalli

Kielimalli sisältää tunnistettavan kielen sääntöihin pohjautuvan todennäköisyyden tunnistettavasta sanasta, yksinkertaisuudessaan se siis ennustaa seuraavan sanan huomioiden edelliset tunnistetut sanat (Saini & Kaur, 2013). Esimerkiksi suomen kielessä sanajärjestys "Käytän korttia maksamiseen" on todennäköisempi kuin "Käytän korttia maksa". Äänisignaalin purkamisen yhtenä edellytyksenä on tietää mitä sanoja voidaan puhua, nämä samat listataan järjestelmän sanakirjaan (Gruhn ym., 2011). Nyrkkisääntönä voidaan pitää, että pienempi sanakirja parantaa tunnistustarkkuutta sillä tunnistettavien sanojen todennäköisyys pienenee (Gruhn ym., 2011).

2.3 Puheentunnistuksen käyttötarkoitukset

Puheen muuttaminen tekstiksi on jo kauan ollut poikkeuksellisen houkutteleva tutkimusaihe sillä se parantaa ihmisten ja tietokoneiden välistä vuorovaikutusta merkittävästi (Khilari & P., 2015). Kuten aikaisemmin mainittiin, automaattisella puheentunnistuksella mahdollistetaan ihmispuheen ymmärtäminen tietokoneen avulla. Ymmärtämisellä tarkoitetaan tässä tapauksessa koneen asianmukaista reagointia sisääntulevaan ja tunnistettuun puheeseen (Saini & Kaur, 2013). Puhe on ihmisille luonnollinen tapa

kommunikoida ja sen saaminen suoraan tietokoneelle tekstiksi « hands freenä » voi parantaa tehokkuutta lukuisissa käyttötarkoituksissa (Khilari & P., 2015).

Yksi automaattisen puheentunnistuksen avainkäyttötarkoituksista on esim. keskustelujen, esitysten, lähetysten ja luentojen tunnistaminen puheasiakirjoiksi joiden puhtaaksikirjoitus voisi muuten olla varsin työlästä ja aikaavievää (Furui ym., 2004). Erityisesti käsistään kiireiset käyttäjät kuten radiologit, lentokoneen pilotit ja postinjakajat hyötyvät puheentunnistuksen mahdollistamasta nopeudesta ja vapaudesta (Chavan & Sable, 2013). Automaattinen puheentunnistus mahdollistaa myös lukuisia muita käyttötarkoituksia nykyaikaisissa tietojärjestelmissä, kuten puheohjausta, sanelua ja puhehakua (Gales & Young, 2008). Joissain käyttötapauksissa voi olla myös nopeampaa käyttää puheentunnistusta jos käyttäjälle käsinkirjoittaminen aiheuttaa merkittävästi häiriöitä, täten puheentunnistus voi auttaa myös liikuntarajoitteisia ihmisiä tekstin tuottamisessa (Chavan & Sable, 2013).

2.4 Puheentunnistusteknologian uudet ratkaisut

Vaikka puheentunnistusteknologia pitkälti pohjautuukin yhä Markovin piilomallia hyödyntäviin ratkaisuihin, on se tehnyt useita teknologisia harppauksia kehityksen suhteen. Keinotekoisiiin neuroverkkoihin pohjautuvat (artificial neural networks ANN) puheentunnistusratkaisut esiteltiin jo 1950-luvulla ilman sen suurempaa menestystä, mutta niitä alettiin tutkia uudelleen 1980-luvulla potentiaalisina tunnistustarkkuutta parantavina teknologioina (Juang & Rabiner, 2005). Suurin osa nykyisistä puheentunnistusratkaisuista on HMM ja Gaussin sekoitusmallin (GMM) hybridejä, joissa GMM määrittää kuinka hyvin jokainen HMM-tila sopii vastaanotettuun akustiseen signaaliin (Hinton ym., 2012). Viime vuosina neuroverkkoihin kohdistunut kehitys on kuitenkin todistanut, että syviin neuroverkkoihin (deep neural networks, DNN) pohjautuvat akustiset mallinnukset voivat ylittää GMM:ien suorituskyvyn, erityisesti suuria sanastoja sisältävissä järjestelmissä (Hinton ym., 2012). DNN sisältää enemmän kuin yhden kerroksen piilotettuja muuttujia sen sisään- ja ulostulojen välillä, tämän takia niillä pystytään tunnistamaan tehokkaammin yhteneväisyyksiä foneemien ja äänisignaalin välillä, HMM:n synnyttämistä tiloista (Hinton ym., 2012).

Vaikka syviin neuroverkkoihin pohjautuvat HMM/GMM hybridijärjestelmät mahdollistavat suuren tunnistustarkkuuden, eivät ne silti ole ongelmattomia. Kyseisissä hybridijärjestelmissä jokainen puheentunnistusjärjestelmän moduuli tulee kouluttaa erikseen, tämä on erittäin aikaa vievää ja työlästä (Zhang ym., 2016). Viime vuosien tekoälyn ja neuroverkkojen äärimmäisen nopean edistymisen myötä tutkijat ovat olleet kiinnostuneita kehittämään ”end-to-end” kokonaisratkaisuja automaattista puheentunnistusta varten. Hybridijärjestelmiin verrattuna, end-to-end järjestelmissä puheentunnistuksen kaikki mo-

duulit oppivat yhdessä, tämä yksinkertaistaa järjestelmän prosessia ja parantaa tunnistustarkkuutta (Zhang ym., 2019).

Vastikään esitetyssä huippumodernissa puheentunnistusratkaisussa yhdistetään monireittinen konvoluutiohermoverkko (Multipath Convolutional Neural Network, MCNN) yhdessä konnektionistisen ajallisen luokituksen (Connectionist Temporal Classification, CTC) kanssa eli helpotettuna MCNN-CTC. Kyseisen sanahirviön tarkoituksena on siis korvata kokonaan HMM-GMM pohjainen puheentunnistus. MCNN on laajennettu malli konvolutionaalisesta neuroverkosta (Convolutional Neural Network, CNN) ja sillä on kyky purkaa puheen yksityiskohtia erittäin tehokkaasti johtaen parantuneeseen tunnistuksen suorituskykyyn. CTC voidaan määritellä kohdetoiminnoksi, jolla on täysi kyky hyödyntää puhesignaaliin liittyviä ajallisia- ja spektrillisiä signaaleita samanaikaisesti, eli yksinkertaistettuna se maksimoi ulostulon oikean todennäköisyyden. MCNN-CTC-pohjainen puheentunnistusratkaisu paransi jopa 12% tunnistustarkkuutta verrattuna perinteisiin HMM-pohjaisiin ratkaisuihin. (Zhang ym., 2019.)

3 TERVEYDENHUOLLON POTILASKERTOMUKSET

Terveydenhuollon kasvava kysyntä ja suuret kustannukset asettavat alati paineita sen jatkuvaan kehittämiseen. Rajoitetut resurssit sekä hoidon vaihtelevuus ovat lisänneet kiinnostusta terveydenhuollon suorituskyvyn ja laadun kehittämiseen sekä sen mittaamiseen useissa maailman maissa (Campbell, Roland & Buetow, 2000). Terveydenhuollon suorituskykyä voidaan arvioida sen kyvyllä toimia välittömästi sitä koskeviin julkisiin ja yksilöiden terveydellisiin tavoitteisiin (Arah ym., 2006). Suorituskykyinen terveydenhuoltojärjestelmä pystyy siis tarjoamaan tehokasta ja tasa-arvoista terveydenhuoltoa kaikille sitä tarvitseville (Arah ym., 2006). Terveydenhuoltoalalla onkin tapahtunut merkittävää kehitystä useissa osa-alueissa, jotta se pystyy tarjoamaan alati parempia ja tehokkaampia palveluita. (Ajami, 2016). Tässä luvussa keskitymme terveydenhuoltojärjestelmiin ja niiden yhteen jatkuvasti kehitettävään osa-alueeseen: potilaskertomusten saneluun.

3.1 Terveydenhuoltojärjestelmä

Campbell ym., (2000) määrittävät terveydenhuollon muodostuvan terveydenhuoltojärjestelmistä sekä niiden sisällä toteutettavista toimista yksilöiden terveyden tai hyvinvoinnin parantamiseksi. On tärkeää erottaa terveydenhuollon rakenne sen tarjoajista ja sen vuorovaikutuksen tuloksena syntyvistä lopputuloksista (Campbell ym., 2000). WHO:n (2020) mukaan terveydenhuolto jakautuu kolmeen sidosryhmään, jotka toimivat keskenään määrittääkseen terveydenhuollon ja sen hallinnon :

- Valtio (valtion ylläpitämät organisaatiot ja virastot keskus- ja aluetasolla)
- Terveyspalveluiden tarjoajat (julkiset ja yksityiset klinisten terveyspalveluiden tarjoajat)
- Kansalaiset (väestön edustajat joista tulee terveydenhuollon käyttäjiä)

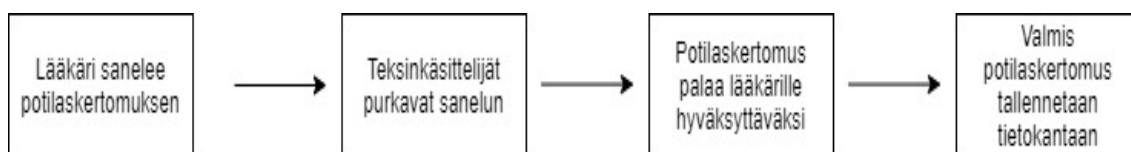
Terveydenhuoltojärjestelmä sisältää siis kaikki toiminnot ja rakenteet, joiden päätarkoituksina on vaikuttaa yksilön terveyteen. Yksilön terveys määräytyy lukuisten toisistaan riippuvien tekijöiden summasta kuten terveydenhuollosta, ympäristöstä, elämäntavasta ja ihmisbiologiasta. Yksilöiden terveyden ja sairauksien välittömän vaikutuksen lisäksi terveydenhuolto voi toimia epäsuorasti myös muihin tekijöihin, jotka vaikuttavat tai ylläpitävät terveyttä. Näitä ovat mm. kansanterveydellisyyteen tähtäävät toimet kuten sairauksien ehkäisy-, edistämisen- ja suojelustrategiat. (Arah ym., 2006.)

Terveydenhuollon käyttömenot suhteessa (%) BKT:hen OECD-maissa ovat olleet lähes poikkeuksetta jatkuvassa nousussa vuodesta 2010-lähtien (OECD Health Statistics, 2019). Vaikka esimerkiksi Suomi nauttii pohjoismaiden pienintä suhdetta terveydenhuollon käyttömenot/asukasluukuun nähden, ovat ne silti suuremmat kuin OECD-mailla keskimäärin (OECD Health Statistics, 2019). Terveydenhuollon suuri osuus BKT:sta, siihen kohdistuvat merkittävät investoinnit ja kasvava osuus työmäärästä aiheuttaa alati paineita terveydenhuoltojärjestelmien kehittämiseen (Folland ym., 2017). Terveydenhuoltojärjestelmiin implementoitu sähköinen terveystietokanta (Electronic Health Record, EHR) onkin ollut kovan kehityksen alla. EHR:llä tarkoitetaan terveystietokantojen systemaattista keräämistä laajoihin, jopa kansallisiin potilastietokantoihin (Liu ym., 2016). EHR-järjestelmien hyötyjä ovat terveystietokantojen nopea saatavuus, terveydenhuollon laadun parantuminen ja vähentyneet terveydenhuoltojärjestelmän kustannukset (Liu ym., 2016). Esimerkiksi Suomessa otettiin käyttöön sähköinen kansallinen terveystietokanta Kanta vaiheittain vuodesta 2010 alkaen. Järjestelmän tarkoituksena on helpottaa kansalaisten asiointia ja parantaa hoidon laatua (Kanta, 2020).

3.2 Potilaskertomukset ja niiden sanelu

Potilasasiakirjoja ja potilaskertomuksia on kerätty potilaista jo siitä asti kun ihmiset osasivat kirjoittaa. Yhdysvalloissa potilaskertomuksia alettiin kerätä systemaattisesti kaikilta potilailta tutkimuksia varten 1900-luvun alussa. (Gillum, 2013). Potilaskertomuksia kirjoitettiin ensiksi käsin josta siirryttiin ajan myötä kirjoituskoneisiin ja nykyisenkaltaisiin tietokoneella tehtäviin potilassaneluihin (Gillum, 2013). Suomessa sosiaali- ja terveysministeriön potilaskirja-asetus (298/2009) määrittää potilaskertomuksen terveydenhuollon ammattihenkilön kirjaamaksi aikajärjestyksessä eteneväksi muuttumattomaksi transkriptioksi. Potilaskertomukseen tehdään merkinnät jokaisesta potilaan palvelutapahtumasta ja siitä tulee käydä ilmi potilaan palvelutapahtumalle merkittävät tiedot kuten, tulosyy, esitiedot, nykytila, havainnot, taudinmääritys, johtopäätökset, hoidon suunnittelu sekä loppulausunto (298/2009).

Tyypillinen digitaalisen sanelujärjestelmän prosessi potilaskertomusten luomiseksi etenee kuvion 3 mukaan:



KUVIO 3 Potilaskertomuksen saneluprosessi

Useimmat potilaskertomukset sanellaan itse lääkärin tai sairaanhoitajan toimesta potilaan vastaanoton jälkeisenä aikana. Saneluun voidaan käyttää mikrofonillista sanelinta tai järjestelmän kanssa yhteensopivaa puhelinta. Potilaskertomuksen merkinnät sanellaan ylhäällä mainitun kaavan mukaisesti potilasasiakirjojen yhteneväisyyden ylläpitämiseksi. Sanelun jälkeen äänitiedostot lähetetään sanelutiedostopalvelimelle odottamaan sihteerien tai tekstinkäsittelijöiden purkamista. Sanelut ohjataan sopiville purkajille jotka kirjoittavat potilaskertomuksen puhtaaksi. Potilaskertomukset lähetetään takaisin keskuspalvelimelle josta lääkäri löytää ne tarkastettavaksi ja hyväksyttäväksi. Lopulta potilastekstit tallennetaan potilastietokantaan arkistointia varten. (Mohr ym., 2003.)

3.3 Perinteisen potilassanelun haasteet

Potilasasiakirjojen dokumentoinnin kasvavan määrän ja digitalisaation myötä potilaskertomukset ovat olleet vauhdikkaan kehityksen alla. Radiologia on yksi tärkeimmistä potilaskertomuksia tuottavista terveydenhuollon yksiköistä, sillä radiologien havainnot ja raportit ovat ratkaisevassa asemassa potilaan hoidossa (Chang, Strahan & Jolley, 2011). On elintärkeää että radiologien lausunot ehtivät lääkärille ajoissa ja tarkasti. Kyseisen prosessin hitaus onkin yksi perinteisen potilassanelun suurimmista haittapuolista (Chang ym., 2011). Potilaskertomusten saneluiden rakenteellinen puute ja viivästyksien prosessin toimintaketjussa voi myös lisätä todennäköisyyttä hoidolle tärkeiden elementtien puutteesta lopullisessa raportissa (Laflamme ym., 2005).

Kasvatut henkilöstökulut ja kysyntä potilasasiakirjoista hoitohenkilökunnan käytössä ovat pakottaneet useat maat ulkoistamaan potilaskertomusten purkupalveluita. Monimutkaiset lääketieteelliset termit, lääkäreiden menettelytapojen tunteminen ja terveydenhoidon vaatimat raportointitavat vaativat jatkuvaa huomiota potilaskertomusten purkajilta. Tekstinkäsittelijöiltä vaaditaan lääketieteellisen terminologian hyvä tuntemus ja potilaskertomusten tulee yleensä saavuttaa 99.8% tarkkuus ja virheettömyys. Ulkoistettujen tekstinkäsittelijöiden äidinkieli voi olla eri kuin sanelevan terveydenhuollon ammattilaisen, eivätkä tekstinkäsittelijät välttämättä tunne lääkäreiden sanelemia ammattitermejä. Tämä voi johtaa huonommin tuotettuihin potilaskertomuksiin, jotka vaativat vielä erillisten tekstinkäsittelijöiden tarkastusta. Pahimmassa tapauksessa virheet voivat

vaarantaa myös potilasturvallisuuden ja johtaa hoitovirheisiin. (Kshetri & Dholakia, 2011.)

Ulkomaille ulkoistetut potilasasiakirjojen tuotantopalvelut voivat aiheuttaa myös vakavia tietoturvahyökkäyksiä. Vuonna 2003 yhdysvaltalaiselle sairaalalle työskentelevä pakistanilainen teksinkäsittelijä uhkasi vuotaa luottamuksellisia potilassaneluita internetiin jos hänen palkkaansa ei korotettaisi. Kyseinen tapahtuma herätti huolta potilasasiakirjojen ulkoistamisesta Yhdysvalloissa. (Kshetri & Dholakia, 2011.)

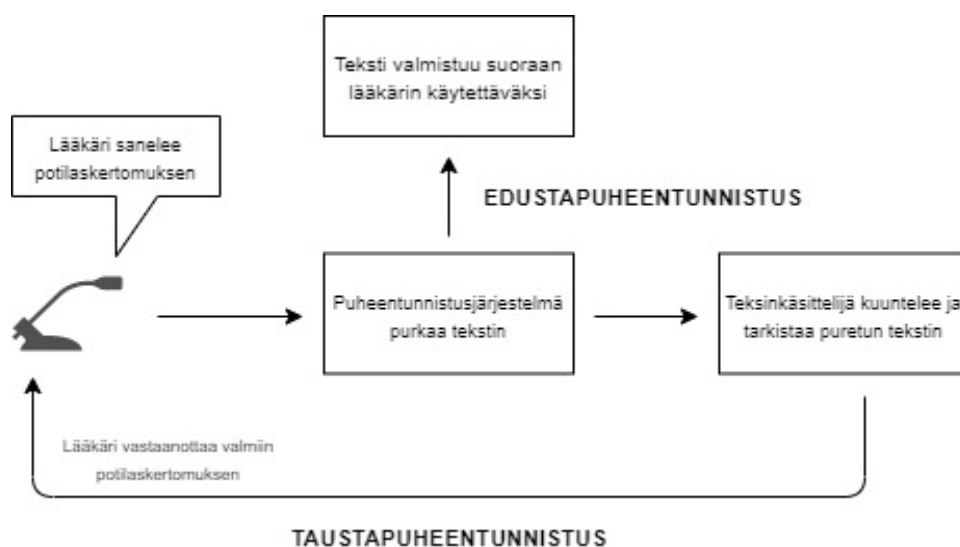
4 PUHEENTUNNISTUS TERVEYDENHUOLLOSSA

Hoitohenkilökunnan loppuunpalamisen ja henkilöstöpulan seurauksena potilasasiakirjojen dokumentointi ja työnkulku on ollut merkittävän kehityksen alla (Saxena ym., 2018). Tämän seurauksena puheentunnistusteknologian käyttöön terveydenhuollossa on panostettu entistä enemmän (Saxena ym., 2018). Puheentunnistus terveydenhuollossa on kuitenkin kaikkea muuta kuin uusi keksintö, sillä ensimmäisiä puheentunnistusjärjestelmiä alettiin käyttää terveydenhuollon dokumentointitarkoituksiin jo 1990-luvun alussa. Tuohon aikaan puheentunnistusjärjestelmät olivat kuitenkin hyvin rajoittuneita ja käyttäjien piti puhua erittäin selkeästi, hitaasti ja pitäen luonnottoman pituisia taukoja (Durling & Lumsden, 2008). Erittäin monimutkaiset alakohtaiset sanastot ja lääketieteellinen terminologia rajoittivat ensimmäiset puheentunnistusjärjestelmät lähinnä radiologien käytettäväksi (Edwards ym., 2017). Teknologian nopea kehittyminen on vauhdittanut ja laajentanut puheentunnistuksen käyttömahdollisuuksia terveydenhuollossa valtavasti viimeisen kymmenen vuoden aikana.

4.1 Puheentunnistus pohjainen potilassanelu

Puheentunnistuksen käyttö terveydenhuollossa on mahdollistanut entistä tehokkaamman ja laadukkaamman potilasasiakirjojen dokumentoinnin viimeisen vuosikymmenen aikana. Kuten jo aikaisemmin mainittiin, käsin tai sanellen tuotetuissa potilaskertomuksissa on merkittäviä rajoitteita. Ne saattavat sisältää joko virheellistä dataa, niiden rakenteellisuus on väärää tai luettavuus on huonoa. Tämä aiheuttaa suoran riskin potilasturvallisuuteen ja pelkästään Yhdysvalloissa 100 000 potilasta menettää vuosittain henkensä hoitotyön dokumentaatiosta johtuvista virheistä. Puheentunnistusteknologian käyttö terveydenhuollon potilassaneluissa johtaa potilasasiakirjojen ja potilasturvallisuuden laadun selvään parantumiseen. (Ajami, 2016).

Terveydenhuollossa käytettävät puheentunnistusjärjestelmät jaetaan yleensä kuvio 4 mukaan kahteen toimintaperiaatteeseen: edustapuheentunnistukseen (front-end speech recognition) ja taustapuheentunnistukseen (back-end speech recognition). Edustapuheentunnistuksessa lääkärin sanelema teksti muodostuu puheentunnistusjärjestelmässä välittömästi valmiiksi ja muokattavaksi tekstiksi. Kyseisen mallin mukaisesti lääkäri voi viimeistellä tekstin ja tallentaa sen potilastietokantaan ilman perinteisen saneluprosessin tuomaa viivettä. Jos edustapuheentunnistus ei ole käytettävissä tai lääkäri ei jostain syystä halua käyttää sitä, voidaan potilaskertomus myös sanella taustapuheentunnistuksena. Kyseinen prosessi ei eroa lääkärin näkökulmasta millään lailla perinteiseen saneluun verrattuna. Saneltu teksti kulkee puheentunnistusjärjestelmän kautta tekstinkäsittelijälle muokattavaksi ja tarkastettavaksi. Tämän jälkeen se palaa normaaliin tapaan lääkärin käytettäväksi ja tallennettavaksi potilastietokantaan. (Vogel ym., 2015).



KUVIO 4 Potilassaneluiden puheentunnistusjärjestelmä

Puheentunnistusteknologia on tehnyt merkittäviä harppauksia viimeisten kymmenen vuoden aikana. Nykyjärjestelmissä on merkittävästi parempi tunnistustarkkuus, suurempi sanakirjasto ja mahdollisuus sanella täysin luonnollisella puherytmillä. Tämä on mahdollistanut puheentunnistusteknologian laajan hyödyntämisen radiologian lisäksi myös muissa terveydenhuollon osa-alueissa kuten patologiassa, avohoidossa, päivystyksessä, ensihoidossa ja kirurgiassa. (Ajami, 2016.) Puheentunnistusta on käytetty pitkään myös Suomen terveydenhuollossa. Oy Konttorityö on maahantuonut Nuance Healthcare teknologiaan perustuvia terveydenhuollon puheentunnistusjärjestelmiä useaan suomalaissairaalaan kuten esimerkiksi HUS-kuvantamiselle, jotka ovat käyttäneet puheentunnistusta jo vuodesta 2004 lähtien (Speech.fi, 2020). Vuonna 2020 alueellisen potilastietojärjestelmä Apotin myötä käyttöön otettiin myös koko HUSin laajui-

nen puheentunnistusratkaisu, jonka käyttöaste on nopeasti noussut jopa 83%: tiin lääkäreiden keskuudessa (Hus.fi, 2020).

Terveydenhuollossa käytettävien puheentunnistusjärjestelmien heikkouksina ovat aina olleet massiiviset lääketieteelliset alakohtaiset sanakirjat, joiden harjoittaminen järjestelmään on ollut haastavaa. Useat puhujasta riippuvat järjestelmät oppivat tehokkaasti käyttäjien itse korjaamista virheistä ja sanojen lisäämisestä sanastoon (Vogel ym., 2015). Korjaavia toimenpiteitä voivat olla esimerkiksi oman ääntämysmallin lisääminen tietyille sanalle tai virheellisesti tunnistettujen sanojen manuaalinen korjaaminen tekstistä (Vogel ym., 2015). Vastaavalla oppivalla järjestelmällä saavutettiin Saksassa 97% tunnistustarkkuus nopeasti puheentunnistusjärjestelmän käyttöönoton jälkeen (Ajami, 2016). Kyseinen prosessi voidaan kuitenkin mieltää sanelijoiden toimesta työlääksi ja jättää täten tekemättä, tämä estää järjestelmän jatkuvan kehittymisen. Vastikään kehitetyt huippumodernit neuroverkkoihin perustuvat end-to-end puheentunnistusjärjestelmät mahdollistavat entistä nopeamman ja tehokkaamman puheentunnistusjärjestelmien kouluttamisen ja käyttöönoton (Chiu ym., 2018). Kyseisissä ”kuuntele, osallistu ja lausu” (Listend, Attend and Spell, LAS) ratkaisuisissa puheentunnistusjärjestelmän kielimalli oppii kaikilta järjestelmän käyttäjiltä samanaikaisesti, mahdollistaen entistä tehokkaamman käyttöönoton (Chiu ym., 2018).

Tämänhetkiset puheentunnistusjärjestelmät tarjoavat useita erilaisia ratkaisuja ja käyttötapoja terveydenhuollossa. Taulokossa 1 vertailemme kahta yleisintä puheentunnistusratkaisua perinteiseen saneluprosessiin.

Puheentunnistus- ratkaisun käyttötapaukset terveydenhuollossa	Ratkaisun hyödyt	Ratkaisun haitat	Lähteet
Edustapuheen- tunnistus	<ul style="list-style-type: none"> - Saneltu teksti on välittömästi lääkärin käytettävissä - Raportin läpimenoaika on teoriassa olematon - Raporttien volyymin kasvu - Käyttäjätyytyväisyyden kasvu - Voi tuottaa merkittävät kustannukselliset säästöt 	<ul style="list-style-type: none"> - Tunnistustarkkuus voi olla pieni - Sanelun virheet tulee korjata itse - Suuri taustamelu estää tunnistuksen - Massiiviset sanastot aiheuttavat haasteita - Järjestelmän käyttöönotto kallista - Virheet mahdollinen uhka potilasturvallisuudelle - Järjestelmän harjoittaminen hidasta 	Vogel ym., (2015), Ajami (2016), Saxena ym., (2018), Edwards ym., (2017), Hodgson & Coiera (2016), Zhang ym., (2016), (Chiu ym., 2018)

Taustapuheentunnistus	<ul style="list-style-type: none"> - Tuottaa tarkkoja ja laadukkaita raportteja - Säästää paljon aikaa raporttien käsittelyssä - Ei eroa lääkäreiden näkökulmasta perinteisestä potilassanelusta - Raporttien volyymin kasvu - Jopa täysin "hands free" 	<ul style="list-style-type: none"> - Järjestelmän käyttöönotto kallista - Massiiviset sanastot aiheuttavat haasteita - Raporttien läpimenoaika hitaampi kuin edustapuheentunnistuksessa - Järjestelmän harjoittaminen hidasta 	Vogel ym., (2015), Edwards ym., (2017), Ajami ym., (2016), Zhang ym., (2016), (Khilari & P., 2015).
Perinteinen potilassanelu	<ul style="list-style-type: none"> - Kaikista suurin tunnistustarkkuus (99.8%) - Purkaminen mahdollista myös todella pitkissä ja taustamelua sisältävissä saneluissa 	<ul style="list-style-type: none"> - Raporttien läpimenoaika hidasta, jopa useita päiviä - Suuret henkilöstökustannukset - Virheellisesti puretut sanelut voivat vaarantaa potilasturvallisuuden 	Johnson., (2014), Vogel ym., (2015), Saxena ym., (2018), Kshetri & Dholakia, (2011), (Chiu ym., 2018).

TAULUKKO 1 Puheentunnistusratkaisujen vertailu

4.2 Puheentunnistusjärjestelmän hyödyt

Puhe on ihmisille luonnollinen ja tehokas kommunikointiväline. Se on ideaalinen tapa tuottaa tekstiä hektisissä tilanteissa joissa käyttäjällä ei ole mahdollisuuksia naputella näppäimistöä, kuten terveydenhuollossa. Puheentunnistusjärjestelmän hyödyt voidaan jaotella kolmeen kategoriaan :

- Tehokkuus
- Laatu
- Kustannustehokkuus.

Puheentunnistusjärjestelmä voi mahdollisesti parantaa potilasdokumentoinnin nopeutta ja tehokkuutta hoitohenkilöstön kohdalla (Ajami, 2016). AHIMA:n (American Healthcare Information Management Association) mukaan jopa 93% puheentunnistusjärjestelmän käyttäjistä sanoivat puheentunnistuksen tuovan hyötyjä potilasakiirjojen dokumentoimiseen (Ajami, 2016). Järjestelmä oli parhaimillaan ja kustannustehokkain silloin kun potilassanelut olivat lyhyitä (Ajami, 2016). M. Vogel ym., (2015) tutkimuksen mukaan taas puheentunnistusjärjestelmän käyttö tehosti hoitohenkilökunnan potilasakiirjojen dokumentointinopeutta 26% :lla. Kyseisen tutkimuksen mukaan puheentunnistusjärjestelmän merkittävin vaikutus ei ehkä ole

kuitenkaan itse dokumentointinopeuden kasvu vaan lisääntynyt dokumentointimäärä mitä sanelijat pystyivät tuotamaan. Perinteisessä saneluprosessissa potilaskertomusten läpimenoaika (Report Turnaround Time, RTT) voi olla useita päiviä (Johnson ym., 2014). Radiologiassa toteutetun tutkimuksen mukaan pelkästään taustapuheentunnistusjärjestelmä vähensi potilaskertomusten läpimenoaikaa jopa 81% :lla, samalla kasvattaen kiireellisten tunnin sisällä vaadittavien raporttien valmistusaikaa 36% :lla (Johnson ym., 2014). Myös muissa tutkimuksissa havaittiin, että puheentunnistuksen käyttö avohoidossa vähensi raporttien läpimenoaikaa jopa 5 päivällä ja samalla lyhentäen potilaan luona tarvittavaa vierailuaikaa 9 :stä minuutista 3 :een minuuttiin. Saxena ym., (2018) tutkimuksen mukaan 84% käyttäjistä tunsivat puheentunnistusjärjestelmän parantaneen ja tehostaneen heidän työnkulkuaan ja saman tutkimuksen jälkikyselyssä 95% käyttäjistä olivat yhtä mieltä siitä, että puheentunnistusjärjestelmän käyttöönotto oli hyvä asia.

Potilaskertomusten tulee olla laadullisesti mahdollisimman eheitä, sillä pienetkin virheet voivat johtaa potilasturvallisuuden vaarantumiseen ja pahimmassa tapauksessa hoitovirheisiin. Tutkimusten mukaan puheentunnistusjärjestelmillä on useissa tapauksissa päästy jopa 98.5% tunnistustarkkuuteen verrattuna ammattitekstinkäsittelijöiden 99.6% :iin (Johnson ym., 2014). Pienemmästä tunnistustarkkuudesta huolimatta, puheentunnistuksen avulla purettu potilaskertomukset olivat laadultaan parempia sillä ne eivät sisältäneet kieliopillisia virheitä verrattuna käsin purettuihin teksteihin (Johnson ym., 2014). Lisäksi potilasturvallisuus kasvaa sillä kopiointivirheitä sattuu puheentunnistuksessa vähemmän (Ajami, 2016). Saxenan ym., (2018) tutkimuksen mukaan 81% käyttäjistä tunsivat puheentunnistuksen parantaneen potilaskertomusten laadullisuutta ja eheyttä. Tämän seurauksena jopa 60% käyttäjistä huomasi käyttäneensä vähemmän aikaa vastatessaan kysymyksiin ja tarkennuksiin joita potilaskertomuksista yleensä jälkikäteen ilmeni (Saxena ym., 2018).

Terveystieteiden huollon kustannusten jatkuva kasvu on johtanut valtaviin investointeihin ympäri maailmaa kustannustehokkuuden parantamiseksi, useat tutkimukset puhuvatkin puheentunnistuksen tuomien säästöjen puolesta. Saxenan ym., (2018) tutkimuksen mukaan puheentunnistusjärjestelmän käyttö toi merkittäviä säästöjä tutkimusaikana. Perinteisten saneluprosessin kustannukset ilman puheentunnistusta olivat 89,4 tuhatta dollaria kuukaudessa, kun taas puheentunnistuspilotin aikana kustannukset tippuivat radikaalit 81.3% eli 16.7 tuhanteen dollariin kuukaudessa (Saxena ym, 2018). Radiologiassa tehdyn tutkimuksen mukaan puheentunnistuksella saavutettu potilaskertomusten läpimenoajan merkittävä lyhentyminen johti 334,000 dollarin vuosittaisiin säästöihin (Johnson ym., 2014). Kaikissa tutkimuksissa ei kuitenkaan päästy yhtä merkittäviin säästöihin. Ajami (2016) mukaan ensihoidossa käytetyllä puheentunnistuksella saavutettiin 22% säästöt verrattuna perinteiseen potilassaneluun niiden lääkäreiden kohdalla jotka järjestelmää käyttivät. Kahdessa yhdysvaltalais tutkimuksessa havaittiin, että puheentunnistuksen käyttö johti 76,250 dollarin vuosittaisiin säästöihin

vähentyneenä tekstinkäsittelijöiden henkilöstökustannuksina, tutkimuksissa ei otettu kuitenkaan huomioon järjestelmän käyttöönotosta syntyviä kustannuksia (Hodgson & Coiera, 2016).

4.3 Puheentunnistusjärjestelmän haittapuolet

Puheentunnistus ei ole kuitenkaan ongelmaton kaikenkattava ratkaisu terveydenhuollon potilasasiakirjojen dokumentointiin. Lääketieteellinen sanelu nimitäin eroaa suuresti normaalista keskustelusta. Äärimmäisen spesifit terminologiat, tuhannet lääkkeet, nopea sanelutahti ja hälyinen ympäristö aiheuttavat merkittäviä haasteita puheentunnistuksen tunnistustarkkuudelle (Edwards ym ym . Vaikka puheentunnistusjärjestelmät ovat kehittyneet erittäin tarkkoiksi viime vuosien aikana, ei mikään järjestelmä ole silti täydellinen. Järjestelmän menestys riippuu sen ohjelmistosta ja mahdollisimman kattavasta sanakirjasta, näistä huolimatta järjestelmät tuottavat virheitä saneluiden aikana jotka voivat olla merkittäviä potilasturvallisuuden kannalta (Ajami, 2016). Hodgson & Coiera (2016) tutkimuksen mukaan puheentunnistusjärjestelmän käyttö kasvat- ti potilaskertomusten virheiden määrää lähes kaikissa testitapauksissa. Useissa tapauksissa järjestelmä tuotti luokan 3 virheitä, jotka terveydenhuollossa tar- koittavat potilasturvallisuuden vaarantavia virheitä. Näitä virheitä olivat mm. virheelliset lääkemääräykset ja puutteet laboratorion tuloksissa (Hodgson & Coiera, 2016). Saxena ym., (2018) kritisoivat kuitenkin tutkimuksessaan Hodgson & Coieran tutkimuksen lyhyttä puheentunnistusjärjestelmän testiaikaa (120 päivää verrattuna Saxenan ym., (2018) 31 kuukauteen).

Tästä huolimatta myös muut tutkimukset kritisoivat puheentunnis- tusjärjestelmien tarkkuutta potilassaneluissa. Suurin osa terveydenhuollossa käytettävistä puheentunnistusjärjestelmistä ovat puhujakohtaisia ja kuten aikai- semmin mainittiin, tämä vaatii käyttäjiltä toimenpiteitä, jotta järjestelmä oppii puhujan henkilökohtaisen äänimallin (Vogel ym., 2015). Järjestelmän tuottamat virheet voivat olla runsaslukuisia ilman käyttäjän alkuharjoittelua tai virheiden jatkuvaa korjausta (Vogel ym., 2015). Vogel ym., (2015) tutkimuksen mukaan useat testikäyttäjät eivät saavuttaneet puheentunnistuksen tuomaa tehokkuutta, sillä heillä kului jatkuvasti aikaa puretun tekstin korjaamiseen. On kuitenkin hyvä huomioida yllä mainituista tutkimuksista se, että Johnson ym., (2014) tut- kimuksen mukaan puheentunnistusjärjestelmien tunnistustarkkuudet vaihtele- vat huomattavasti eri palveluntarjoajien kesken.

Puheentunnistusjärjestelmien kuten kaikkien tietojärjestelmien me- nestys riippuu kuitenkin siitä näkevätkö käyttäjät järjestelmän hyödylliseksi. Ajami (2016) tutkimuksen mukaan hoitohenkilökunnan muutosvastarinta ai- heuttaa suuria haasteita puheentunnistusjärjestelmien käyttöönotolle. Järjestel- män tuottamiin virheisiin turhautuneet käyttäjät palasivat perinteiseen potilas- kertomusten kirjaamiseen, vaikka sitkeä käyttö olisikin kehittänyt järjestelmää tarkemmaksi (Ajami, 2016). Erityisesti korkean työtaakan yksiköissä kuten päi- vystyksessä työskentelevät lääkärit eivät nähneet puheentunnistuksen tuovan

hyötyjä heidän työnkulkunsa hallintaan (Ajami, 2016). Käyttäjille esitetty väärä informaatio, hidas oppimiskäyrä, rajoitettu saatavuus sekä suuret kustannukset voivat luoda haasteita puheentunnistusjärjestelmän käyttöönottoon (Vogel ym., 2015).

Useissa tutkimuksissa osoitetaan puheentunnistuksen kustannustehokkuus, mutta harvassa mainitaan sen synnyttämiä kuluja. Ajami (2016) mukaan harva huomioi puheentunnistuksen käyttöönoton mittavia kustannuksia. Niitä syntyy kehiteltävistä ja käytettävistä ohjelmistoista, laitteistosta kuten mikrofoneista, koulutuskuluista sekä siihen kuuluvasta työajasta. Yhdessä käytötutkimuksessa huomattiin puheentunnistuksen käytön johtavan jopa useampiin virheisiin ja suurempiin kuluihin verrattuna perinteiseen potilassaneluun. (Ajami, 2016.)

5 YHTEENVETO

Tässä kandidaatintutkielmassa käytiin läpi puheentunnistuksen teknologioita sekä terveydenhuollon järjestelmiä ja potilassaneluita. Tutkielman tarkoituksena oli tutkia puheentunnistusteknologian käyttöä ja sen saavuttamia mahdollisia hyötyjä terveydenhuollon potilassaneluiden työkaluna. Tutkielmassa vastattiin kysymyksiin:

- *Miten puheentunnistusteknologiaa hyödynnetään terveydenhuollon potilaskertomuksissa?*
- *Mitä hyötyjä puheentunnistusteknologian käyttö voi saavuttaa terveydenhuollon potilaskertomuksissa?*

Kirjallisuuskatsauksena toteutettu tutkielma aloitettiin tarkastelemalla puheentunnistuksen historiaa ja teknologiaa. Puheentunnistuksella tarkoitetaan teknologiaa, jonka avulla tietokone purkaa ihmisen tuottaman puheen tekstiksi (Chavan & Sable, 2013). Purettua tekstiä voidaan tämän jälkeen käyttää lukuisissa käyttötarkoituksissa kuten automaattisina tekstityksinä, puheohjauksena tai tämän tutkielman tarkastelemissa terveydenhuollon potilaskertomusten saneluissa (Chavan & Sable, 2013). Lähes kaikissa nykyisissä puheentunnistusjärjestelmissä käytetään lähinnä Markovin piilomalleihin pohjautuvia algoritmeja purkamisen pohjana (Gales & Young, 2008). Tutkielmassa käytiin läpi myös puheentunnistuksen jatkokehitystä ja havaittiin, että puheentunnistuksen tunnistustarkkuuden odotetaan kehittyvän vielä entisestään tulevaisuudessa kiitos huippumodernien MCNN-CTC-pohjaisten tunnistusteknologioiden.

Kirjallisuuskatsauksessa tutkittiin myös terveydenhuoltoa ja sen järjestelmiä. Campbell ym., (2000) määrittävät terveydenhuollon muodostuvan terveydenhuoltojärjestelmistä sekä niiden tuottamista toimista yksilöiden terveyden tai hyvinvoinnin parantamiseksi. Tutkielmassa havaittiin, että terveydenhuollon kustannusten jatkuva kasvu OECD-maissa on johtanut massiivisiin investointeihin sen kustannustehokkuuden ja laadun parantamiseksi. Kyseisen kappaleen tarkoituksena on avata motivaatiota sille, miksi puheentunnistuksen

kaltaisia järjestelmiä sovelletaan terveydenhuollon potilaskertomusten saneluissa.

Puheentunnistuksen käyttö potilaskertomusten saneluissa ei ole uutta, mutta sen tarjoamaa potentiaalia on alettu hyödyntämään vasta viimeisen vuosikymmenen aikana tapahtuneen teknisen kehityksen johdosta. Tutkimustuloksina todettiin, että puheentunnistuksen käyttö potilaskertomusten luomisessa tuo selviä hyötyjä verrattuna perinteiseen saneluprosessiin. Lähes jokaisessa tutkimuksessa havaittiin, että puheentunnistuksen käyttö nopeuttaa potilaskertomusten läpimenoaikaa huomattavasti. Useat tutkimukset puhuivat myös puheentunnistuksen tuomista kustannussäästöistä ja käyttäjätyytyväisyyden kasvusta. Vastaavasti tutkielmassa todettiin puheentunnistuksen tuomat ristiriidat. Puheentunnistuksen tunnistustarkkuus vaihtelee merkittävästi palveluntarjoajan mukaan. Järjestelmän tuottamat virheet ovat yleisiä ja ne aiheuttavat selvän riskin potilasturvallisuudelle. Puheentunnistuksen potentiaali on kiistaton. Se on yksi parhaista ja tehokkaimmista tavoista tuottaa paljon tekstiä nopealla tahdilla eli juuri sitä mitä terveydenhuollon dokumentointitarpeissa kaivataan. Tästä huolimatta puheentunnistus vaatii vielä paljon kehitettävää ennen kuin se saavuttaa potentiaalinsa terveydenhuollon dokumentoinnin työvälineenä.

Tutkielman tavoitteena oli tarjota mahdollisimman ajantasainen vastaus tutkimuskysymyksiin. Tekoälyn, neuroverkkojen ja puheentunnistusteknologian äärimmäisen nopea kehittyminen loivat haasteita tarpeeksi ajankohtaisten aineistojen etsintään. Puheentunnistuksen hyödyntäminen terveydenhuollossa vaihtelee merkittävästi maittain ja jopa sairaaloittain. Tämän seurauksena tutkielman suurimpana rajoitteena oli itse puheentunnistusteknologian käyttöä terveydenhuollossa käsittelevän aineiston suppea määrä. Onneksi löydetyn aineiston laatu oli kuitenkin pääosin hyvä ja niitä oli vertaisarvioitu esimerkiksi terveydenhuollon alan julkaisuissa. Aineistoihin tulee kuitenkin suhtautua kriittisesti sillä joidenkin tutkimusten mittakaavat ja tutkimustavat vaihtelivat huomattavasti. Aineistojen etsinnässä huomattiin, että puheentunnistuksen synnyttämiä todellisia kokonaiskustannuksia terveydenhuollossa tutkittiin vähän. Useissa tutkimuksissa todettiin puheentunnistuksen tuovan suuria säästöjä ottamatta huomioon järjestelmän kehityksen, käyttöönoton, koulutuksen ja ylläpidon hintaa. Jatkotutkimusaiheina voisikin olla puheentunnistuksen käytön todellinen kustannustehokkuus terveydenhuollossa. Hyödyllisiä tutkimusaiheita voisivat myös olla puheentunnistusjärjestelmien käytön hyväksyminen hoitohenkilökunnan puolesta ja tunnistustarkkuus huippumoderneilla tunnistusteknologioilla.

LÄHTEET

- Ajami, S. (2016). Use of speech-to-text technology for documentation by healthcare providers. *National Medical Journal of India*, 29(3), 148–152.
- Arah, O. A., Westert, G. P., Hurst, J., & Klazinga, N. S. (2006). A conceptual framework for the OECD Health Care Quality Indicators Project. *International Journal for Quality in Health Care*, 18(SUPPL. 1), 5–13.
- Campbell, S. M., Roland, M. O., & Buetow, S. A. (2000). Defining quality of care. *Social Science and Medicine*.
- Chang, C. A., Strahan, R., & Jolley, D. (2011). Non-clinical errors using voice recognition dictation software for radiology reports: A retrospective audit. *Journal of Digital Imaging*, 24(4), 724–728.
- Chavan, R. S., & Sable, G. S. (2013). An overview of speech recognition using HMM. *International Journal of Computer Science and Mobile Computing*, 2(6), 233-238.
- Chiu, C. C., Tripathi, A., Chou, K., Co, C., Jaitly, N., Jaunzeikare, D., Kannan, A., Nguyen, P., Sak, H., Sankar, A., Tansuwan, J., Wan, N., Wu, Y., & Zhang, X. (2018). Speech recognition for medical conversations. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2018-Septe*, 2972–2976.
- Durling, S., & Lumsden, J. (2008). Speech recognition use in healthcare applications. *MoMM2008 - The 6th International Conference on Advances in Mobile Computing and Multimedia*, 473–478.
- Edwards, E., Salloum, W., Finley, G. P., Fone, J., Cardiff, G., Miller, M., & Suendermann-Oeft, D. (2017). Medical speech recognition: Reaching parity with humans. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10458 LNAI, 512–524.
- Folland, S., Goodman, A. C., & Stano, M. (2017). The Economics of Health and Health Care. In *The Economics of Health and Health Care*.
- Furui, S., Kikuchi, T., Shinnaka, Y., & Hori, C. (2004). Speech-to-Text and Speech-to-Speech Summarization. *Transactions on Speech and Audio Processing*, 12(4), 401–408.

- Gales, M., & Young, S. (2008). The application of hidden Markov models in speech recognition. *Foundations and Trends® in Signal Processing*, 1(3), 195-304.
- Gillum, R. F. (2013). From papyrus to the electronic tablet: A brief history of the clinical medical record with lessons for the digital age. *American Journal of Medicine*, 126(10), 853–857.
- Gruhn, R. E., Minker, W., & Nakamura, S. (2011). Statistical pronunciation modeling for non-native speech processing. Springer Science & Business Media, s. 5-17.
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A. R., Jaitly, N., ... & Sainath, T. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal processing magazine*, 29.
- Hodgson, T., & Coiera, E. (2016). Risks and benefits of speech recognition for clinical documentation: A systematic review. *Journal of the American Medical Informatics Association*, 23(e1), e69–e179.
- HUS (2020). Haettu osoitteesta <https://www.hus.fi/hus-tietoa/uutishuone/Sivut/Apotti-toiminut-kokonaisuutena-odotetusti.aspx> 4.3.2020.
- Johnson, M., Lapkin, S., Long, V., Sanchez, P., Suominen, H., Basilakis, J., & Dawson, L. (2014). A systematic review of speech recognition technology in health care. *BMC Medical Informatics and Decision Making*, 14(1).
- Ju, Z., Liu, Y., Zhou, D., Vi, P., & Goebel, R. (2019). Series Editors.
- Juang, B. H., & Rabiner, L. R. (2005). Automatic speech recognition—a brief history of the technology development. Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara, 1, 67.
- Kanta (2020). Haettu osoitteesta <https://www.kanta.fi/jarjestelmakkehittajat/mita-kanta-palvelut-ovat> 14.2.2020
- Khilari, P., & P., B. V. (2015). A Review on Speech To Text Conversion Methods. *International Journal of Advanced Research in Computer Engineering & Technology*, 4(7), 3067–3072.
- Kruse, C. S., Kristof, C., Jones, B., Mitchell, E., & Martinez, A. (2016). Barriers to Electronic Health Record Adoption: a Systematic Literature Review.

Journal of Medical Systems, 40(12).

Kshetri, N., & Dholakia, N. (2011). Offshoring of healthcare services: The case of US-India trade in medical transcription services. *Journal of Health, Organisation and Management*, 25(1), 94-107.

Laflamme, M. R., Dexter, P. R., Graham, M. F., Hui, S. L., & McDonald, C. J. (2005). Efficiency, comprehensiveness and cost-effectiveness when comparing dictation and electronic templates for operative reports. *AMIA ... Annual Symposium Proceedings / AMIA Symposium*. AMIA Symposium, 425-429.

Liu, Z., Weng, J., Li, J., Yang, J., Fu, C., & Jia, C. (2016). Cloud-based electronic health record system supporting fuzzy keyword search. *Soft Computing*, 20(8), 3243-3255.

Mohr, D. N., Turner, D. W., Pond, G. R., Kamath, J. S., De Vos, C. B., & Carpenter, P. C. (2003). Speech recognition as a transcription aid: A randomized comparison with standard transcription. *Journal of the American Medical Informatics Association*, 10(1), 85-93.

OECD Health Statistics 2019 - July 2019. Haettu osoitteesta https://www.kuntaliitto.fi/sites/default/files/media/file/OECD%20Health%20Statistics%202019_Taloustietoja_taulukot.pdf 12.2.2020

Saini, P., & Kaur, P. (2013). Automatic speech recognition: A Review. *ICEIS 2003 - Proceedings of the 5th International Conference on Enterprise Information Systems*, 1(iii), IS5-IS10.

Saxena, K., Diamond, R., Conant, R. F., Mitchell, T. H., Gallopyn, I. G., & Yakimow, K. E. (2018). Provider Adoption of Speech Recognition and its Impact on Satisfaction, Documentation Quality, Efficiency, and Cost in an Inpatient EHR. *AMIA Joint Summits on Translational Science Proceedings*. AMIA Joint Summits on Translational Science, 2017, 186-195.

Sosiaali- ja terveystieteiden ministeriön asetus potilasasiakirjoista 298/2009. Haettu osoitteesta <https://www.finlex.fi/fi/laki/ajantasa/2009/20090298> 11.2.2020.

Vogel, M., Kaisers, W., Wassmuth, R., & Mayatepek, E. (2015). Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial. *Journal of Medical Internet Research*, 17(11), 1-12.

WHO World Health Organization. (2020). Haettu osoitteesta

https://www.who.int/health-topics/health-systems-governance#tab=tab_1

Zhang, W., Zhai, M., Huang, Z., Liu, C., Li, W., & Cao, Y. (2019, August). Towards End-to-End Speech Recognition with Deep Multipath Convolutional Neural Networks. In International Conference on Intelligent Robotics and Applications (pp. 332-341). Springer, Cham.

Zhang, Y., Pezeshki, M., Brakel, P., Zhang, S., Laurent, C., Bengio, Y., & Courville, A. (2016). Towards end-to-end speech recognition with deep convolutional neural networks. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 08-12-Sept, 410-414.