

Data-analytiikan ja teknologian hyödyntäminen huippu-urheilussa

Kai Lehtinen

Haaga-Helia ammattikorkeakoulu

Amk-opinnäytetyö

2021

Tradenomin tutkinto

Tiivistelmä

Tekijä(t) Kai Lehtinen
Tutkinto Tradenomi
Raportin/Opinnäytetyön nimi Data-analytiikan ja teknologian hyödyntäminen huippu-urheilussa
Sivu- ja liitesivumäärä 40 + 0
<p>Tässä tutkielmassa vertaillaan ja kuvataan erilaisten tieteisartikkeleiden, kirjojen ja opinnäytetöiden pohjalta, millaisia datankeräysmenetelmiä, data-analytiikkaa, tekoälyn ja teknologian keinoja erilaisissa urheilulajeissa käytetään ja miten niitä hyödynnetään. Tutkielma toteutetaan kuvailevana kirjallisuuskatsauksena, jossa pyritään tunnistamaan, vahvistamaan tai kyseenalaistamaan aikaisempien tutkimusten esiin nostamia kysymyksiä datan ja teknologian hyödyntämisestä huippu-urheilussa.</p> <p>Tutkielma tarkastelee erityisesti mitä dataa pelaajista, urheilijoista ja joukkueista kerätään erilaisten teknologioiden avulla. Tutkielma rajautuu dataan ja teknologialaitteisiin, mitä hyödynnetään urheilusuoritusten parantamiseen. Tutkielma ei tutki dataa urheilun liiketoiminnan näkökulmasta.</p> <p>Tutkielman tavoitteena on luoda eheä näkemys datankeräämisestä, data-analytiikasta ja teknologian hyödyntämisestä huippu-urheilussa. Tutkielma vastaa seuraaviin kysymyksiin: minkälaisia datankeräysmenetelmiä, data-analytiikkaa ja teknologiaa käytetään erilaisissa urheilulajeissa? Miten urheilussa kerätty data vaikuttaa urheilijoihin, joukkueisiin ja itse urheiluun? Mihin urheilussa kerättyä dataa käytetään?</p> <p>Aineistoa aiheesta löytyi runsaasti, mutta suurin osa aineistosta ei läpäissyt aineistokriteereitä, jotka rajattiin tieteellisesti valideihin lähteisiin. Aihetta ei ole tutkittu tai vertailtu erilaisten urheilulajien näkökulmasta, vaan useat tutkimukset rajautuivat vain yhteen tiettyyn urheilulajiin.</p> <p>Tutkielman lopussa on tulokset sekä kirjoittajan henkilökohtaista pohdintaa tuloksista. Tutkielmasta selviää, että dataa on alettu hyödyntämään paljon eri urheilulajeissa ja tulokset olivatkin hyvin samanlaisia urheilulajista riippumatta. Tämä kuvaileva kirjallisuuskatsaus kokosi yhteen merkittävimmät löydökset eri lajien väliltä.</p>
Asiasanat Data, big data, data-analytiikka, IoT, urheiluanalytiikka ja tekoäly.

Sisällys

1	Johdanto.....	1
2	Opinnäytetyön toteutus	5
2.1	Kuvaileva kirjallisuuskatsaus	5
2.2	Aineiston keruumenetelmät	5
3	Teoreettinen tausta ja käsitteet.....	7
3.1	Big Data	7
3.2	Data-analytiikka	8
3.3	Urheiluanalytiikka.....	9
3.4	Liikuntateknologialaitteet	9
3.5	Tekoäly.....	11
3.6	IoT ja IoT-laitteet.....	12
4	Data-analytiikan ja teknologian hyödyntäminen huippu-urheilussa	14
4.1	Suorituskyvyn mittaaminen joukkueurheilussa KPI-mittareiden avulla	14
4.2	Data-analytiikan ja teknologian hyödyntäminen Liigassa	15
4.3	Data-analytiikan ja teknologian hyödyntäminen NHL:ssä.....	18
4.4	Data-analytiikan ja teknologian hyödyntäminen jalkapallossa.....	21
4.5	Data-analytiikan ja teknologian hyödyntäminen koripallossa	23
4.6	Data-analytiikan ja teknologian hyödyntäminen NFL:ssä	29
5	Pohdinta	33
5.1	Yhteenveto.....	33
5.2	Tutkimuksen eettisyys sekä luotettavuus.....	35
5.3	Mahdolliset jatkotutkimukset.....	36
	Lähteet	38

1 Johdanto

Niin urheilu kuin teknologiakin ovat saavuttaneet viime vuosikymmeninä massiiviset mittasuhteet. Urheilu on universaalia, ja sen harrastamiseen on lähes loputon määrä mahdollisuuksia; pikajuoksusta e-urheiluun tai yksilölajeista joukkueurheiluun. Tilastointia urheilusta on tehty pitkään, mutta teknologia on mahdollistanut tilastoinnin kehityksen urheilumaailmassa hyödynnettäväksi urheiludataksi, joka on suunnattu urheilufaneille, joukkueille, urheilijoille itselleen, vedonlyöntiyrityksille ja organisaatioille, jotka harjoittavat urheilua. Teknologian ja datan hyödyntäminen urheilussa on kehittynyt viime vuosina paljon, ja siksi tämä kuvaileva kirjallisuuskatsaus on ajankohtainen hahmottaakseen datan ja data-analytiikan merkityksen nykyisissä suurimmissa urheilulajeissa ja -organisaatioissa.

Tässä opinnäytetyössä tutkitaan verkkojulkaisujen, opinnäytetöiden ja kirjallisuuden pohjalta minkälaisia datan keräysmenetelmiä, data-analytiikan malleja, tekoälyä ja teknologiaa eri urheilulajeissa käytetään ja miten niitä hyödynnetään. Urheilusuoritusten ja urheilijoiden analysoinnin tehtävänä on ymmärtää, miten eri tekijät vaikuttavat menestykseen pelikentillä tai urheiluradoilla. Urheilusuoritusten analysointi alkaa aina mittaamisesta, eli datankeräämisestä. Kerätyn datan avulla voidaan mitata monia erilaisia ilmiöitä, joilla on vaikutusta urheilussa menestymiselle, kuten urheilijan fyysistä kuntoa, biofysiikkaa, terveyttä, kuntoilua ja kuntouttamista. Muita urheilijassa ja urheilussa mitattavia asioita datan avulla ovat myös nopeus, voima, liikkuvuus ja ketteruus. Urheiluanalytiikka ja datankerääminen ovat yleistyneet niin, että dataa kerätään urheilijoista niin harjoituksissa, peleissä kuin vapaa-ajallakin. (Miller 2015, 6.)

Urheiluanalytiikka on kilpailuedun lähde (Miller 2015, 1). Tämä tutkielma ei ota kantaa urheilun liiketoiminnan näkökulmaan vaan tutkielma kartoittaa aiempien tutkimusten perusteella datankeräysmenetelmiä, data-analytiikan keinoja, tekoälyn ja teknologian hyödyntämistä pelaajien, valmentajien, joukkueiden ja urheilun näkökulmasta. Tutkimus on kohdennettu niille, jotka haluavat tietoa urheiluanalytiikasta ja siitä, miten sitä hyödynnetään erilaisissa urheilulajeissa.

Tutkimuksen tavoitteena on löytää eroavaisuuksia ja yhtäläisyyksiä datankeräämismenetelmistä, data-analytiikan käytöstä sekä teknologian ja tekoälyn hyödyntämisestä erilaisissa urheilulajeissa. Tutkielma vastaa kysymyksiin:

1. Minkälaisia datankeräysmenetelmiä, data-analytiikkaa ja teknologiaa käytetään erilaisissa urheilulajeissa?
2. Miten urheilussa kerätty data vaikuttaa urheilijoihin, joukkueisiin ja itse urheiluun?
3. Mihin urheilussa kerättyä dataa käytetään?

Tutkielma etenee tutkielman toteutustavasta aineiston rajaukseen sekä keinoihin, joilla aineisto kerätään. Tämän jälkeen tutkielma siirtyy käsitteiden määrittelyyn, jonka jälkeen se etenee aikaisempien tutkimusten tarkasteluun. Aikaisempien tutkimusten tarkastelun jälkeen esitellään tutkimustulokset, sekä datan ja teknologian käyttöön liittyvät mahdolliset riskit ja ongelmakohdat, jos niitä on. Tutkimus päättyy pohdintaan, josta löytyy mahdolliset jatkotutkimusten aiheet ja omia näkemyksiäni entisenä ammattuurheilijana data-analytiikan ja teknologian hyödyntämisestä huippu-urheilussa.

Tutkielmassa usein käytettyjen lyhenteiden selitykset:

NHL	National Hockey League (NHL) on Yhdysvalloissa ja Kanadassa pelattava jääkiekon miesten ammattilaissarja, jossa pelaa 32 joukkuetta, joista seitsemän on Kanadasta.
NBA	National Basketball Association (NBA) on Pohjois-Amerikassa pelattava miesten koripallon ammattilaissarja, jossa pelaa 30 joukkuetta, joista yksi on Kanadasta.
NFL	National Football League (NFL) on Yhdysvalloissa pelattava amerikkalaisen jalkapallon miesten ammattilaissarja, jossa pelaa 32 joukkuetta.
GPS	Global Positioning System (GPS, suom. maailmanlaajuinen paikallistamisjärjestelmä). GPS-laitteet lukevat satelliittien lähettämää mikroaaltosäteilyä. GPS-laitteiden avulla esimerkiksi urheilujoukkueet voivat paikantaa pelaajien liikkeitä otteluiden aikana.
KPI	Key Performance Indicators (KPI), eli tärkeimpiä suorituskykyyn vaikuttavia tekijöitä. KPI-mittareiden avulla on mahdollista seurata jonkin määritetyn toiminnan tehokkuutta peilaamalla tuloksia asetettuihin KPI-mittareihin.
DM	Data mining (DM) tunnetaan tiedonlouhintana. Tämä tarkoittaa niitä menetelmiä, joilla pyritään löytämään oleellista tietoa eli dataa suurista data massoista.
ML	Machine learning (ML) tunnetaan koneoppimisena, joka on yksi tekoälyn osa-alue. Sen tarkoituksena on saada ohjelmisto

toimimaan paremmin tai jopa itsenäisesti, tarkkaan määriteltyjen tietojen pohjalta.

IoT	Internet of Things (IoT) tarkoittaa esineiden internetiä. IoT perustuu teknisiin laitteisiin, jotka suorittavat automaattista tiedonsiirtoa internet verkkoyhteyden avulla.
AI	Artificial Intelligence (AI) eli tekoäly, on tietokone tai tietokoneohjelma, joka kykenee älykkääseen toimintaan.
SAP	SAP on saksalainen teknologia yritys, joka on Euroopan suurin ja maailman kolmanneksi suurin ohjelmistovalmistaja.

Urheiluun liittyvät keskeiset ammattikäsitteet:

Plus-miinus-tilasto

Plus-miinus-tilasto (+/- tilasto) tunnetaan urheilussa myös nimellä tehotilasto. Sillä tarkoitetaan tehtyjen ja päästettyjen maalien erotusta pelaajan ollessa kentällä tasakentällisin. Tällä tilastolla pystytään tarkastelemaan esimerkiksi pelaajan arvoa hyökkäyspeliin ja puolustuspeliin.

SportVU

Kamerajärjestelmä, joka pystyy kerätä dataa 25 datapistettä sekunnissa. Järjestelmän avulla urheilujoukkueet keräävät dataa pelaajista ja pelivälineestä kentältä. Pelaajien peliasuihin tai suojuksiin kiinnitetään sirut, joiden ansiosta järjestelmä tunnistaa heidät kentältä. Samanlainen siru on myös kiinnitettyinä pelivälineeseen.

Taktiikka

Taktiikalla tarkoitetaan urheilussa joukkueen tietoisesti rakennettua ja sovittua tapaa pelata. Taktiikoita on usein useita eri tilanteisiin ja eri kentän alueisiin. Taktiikat vaihtelevat myös urheilulajin mukaan, toiset lajit ovat huomattavasti taktisempia kuin toiset.

Urheiluanalytiikka

Urheiluanalytiikka käsittelee kokoelmaa aiemmasta ja nykyisestä urheiludatasta sekä sen keräämisestä ja analysoimisesta.

Urheiluanalytiikka on myös osa tänä päivänä tapahtuvaa urheilutilastointia.

Puettavat teknologialaitteet

Puettavat teknologia laitteet ovat IoT-laitteita, joilla pystytään seuraamaan esimerkiksi ihmisen sydämensykeä, sijaintia ja liikkumisnopeutta. Tällaisia laitteita ovat esimerkiksi sykemittarit, sykevyöt, älysormukset tai sykeliivit. Urheilussa käytetään usein liivejä tai rinnanympärillä meneviä vöitä.

Datan kerääminen

Datan kerääminen tarkoittaa tiedon keräämistä. Urheilussa dataa kerätään pelaajista esimerkiksi edellä mainituilla puettavilla teknologialaitteilla. Pääsääntöisesti urheilussa datan keräämisellä tarkoitetaan urheilijoista kerättäviä tietoja, tai pelistä kerättäviä tietoja erilaisista pelitapahtumista.

2 Opinnäytetyön toteutus

2.1 Kuvaileva kirjallisuuskatsaus

Tutkielmassa vertaillaan ja kuvataan erilaisten tieteisartikkeleiden, kirjojen ja opinnäytetöiden pohjalta, millaisia datankeräysmenetelmiä, data-analytiikan, tekoälyn ja teknologian keinoja erilaisissa urheilulajeissa käytetään ja miten niitä hyödynnetään. Aikaisemmista tutkimuksista pyritään tunnistamaan ja kyseenalaistamaan aikaisempien tutkimusten esiin nostamia kysymyksiä sekä löytämään mahdollisia eroavaisuuksia ja yhtäläisyyksiä tutkimusten väliltä. Tutkimuksessa perehdytään erityisesti teknologian rooliin urheilussa ja millaisia teknologialaitteita datankeräämisessä ja data-analytiikassa hyödynnetään. Tutkielmassa data-analytiikan ja teknologian keinoilla tarkoitetaan sitä dataa mitä pelaajista, urheilijoista tai joukkueista saadaan teknologialaiteiden avulla heidän harjoittelunsa tai urheilusuorituksen aikana.

Tämä laadullista tutkimussuuntaa edustava opinnäytetyö toteutetaan kuvailevana kirjallisuuskatsauksena. Kuvailevassa kirjallisuuskatsauksessa pyritään tunnistamaan, vahvistamaan tai kyseenalaistamaan aikaisempien tutkimusten esiin nostamia kysymyksiä. Kuvailevassa kirjallisuuskatsauksessa pyritään myös tunnistamaan aikaisempien tietojen ristiriitoja sekä aikaisemman tiedon jättämiä kysymyksiä. (Ahonen ym. 2013, 294.) Uusi tai erilainen näkökulma tulee löytää kuvailevan kirjallisuuskatsauksen avulla. Tämän tutkielman tavoitteena on kohdentua käsitteellisen ja teoreettisen kehyksen rakentamiseen. Kuvaileva kirjallisuuskatsaus soveltuu erityisen hyvin hajanaisiin tai pirstaleisiin aiheisiin (Ahonen ym. 2013, 295). Tämä tutkielma päättyi kuvailevaan kirjallisuuskatsaukseen, sillä kuvailevan kirjallisuuskatsauksen erikoispiirteet tukivat parhaalla mahdollisella tavalla tämän opinnäytetyön tavoitteita sekä lähestymistapaa tutkittavaan aiheeseen. Urheilulajeja sekä niissä käytettyä teknologiaa on runsaasti, joka on luonut hajanaisen lähdeaineistokentän, josta lopulta on tarkoituksena koota onnistunut sekä eheä katsaus datan, että teknologian suhteesta huippu-urheiluun.

2.2 Aineiston keruumenetelmät

Aineiston valinta tutkimukseen tehdään elektronisista tieteellisistä tietokannoista, kuten Google.scholar, HH-Finna ja Helka-kirjasto. Aineistoa valittaessa otetaan huomioon tutkimuskysymys ja se, että aineisto vastaa tutkielman tutkimuskysymykseen. Aineiston valinta tutkielmaan on eksplisiittinen, eli aineiston haut kuvailevaan kirjallisuuskatsaukseen tehdään systemaattisesti. Tässä tapauksessa aineiston eksplisiittisellä hauella tarkoitetaan sitä, että aineistoa haetaan manuaalisesti sähköisistä tietokannoista ja hauissa hyödynnetään aineiston aika- ja kielirajauksia (Ahonen ym.

2013, 296). Aineiston aikarajaukset otetaan huomioon aineistoa kerätessä tarkasti. Aineisto ei voi olla monia vuosia vanhaa, koska teknologian kehitys on mennyt paljon eteenpäin urheilussa viime vuosina. Tutkielman aineiston kielirajaukset rajoittuvat suomeen ja Englantiin. Kuvailevalle kirjallisuuskatsaukselle on tavallista, että aineiston valinnassa on jatkuvaa ja vastavuoroista reflektointia suhteessa tutkimuskysymykseen, jolloin sekä tutkimuskysymys, että siihen vastaamaan valittu aineisto tarkentuvat koko tutkielman ajan (Ahonen ym. 2013, 296). Aineisto tullaan keräämään niin, että aineiston sisäinen vertailu on mahdollista. Aineistosta analysoidaan löytyvän tiedon vahvuuksia ja heikkouksia koko tutkielman ajan, jotta laajempien päätelmien tekeminen aineistosta olisi mahdollista. Tutkielman tavoitteena on onnistuneen kokonaisuuden luominen, eli valitusta aineistosta haetaan tutkielman kannalta merkityksellisiä seikkoja ja ilmiöitä, jotka lopulta kootaan sisällöllisesti kokonaisuuksiksi. (Ahonen ym. 2013, 297.)

3 Teorettinen tausta ja käsitteet

Opinnäytetyön aihe datankerääminen, data-analytiikka ja teknologian hyödyntäminen huippu-urheilussa on nyky-yhteiskunnassa aiheena ajankohtainen. Opinnäytetyön aiheen motiivina oli kiinnostukseni urheiluun sekä edesmennyt oma ammattilaisurani jääkiekkoilijana. Aineistoa löytyy tutkittavasta asiasta runsaasti, mutta suurin osa aineistosta ei läpäissyt aineistokriteereitä, jotka rajattiin tieteellisesti valideihin lähteisiin. Dataa on alettu käyttämään urheilussa yhä enemmän, mutta tutkimuksen aiheena datan hyödyntäminen urheilussa on selvästi vasta kehittymässä ja sitä tutkittaneen tulevaisuudessa lisää. Motiivi tämän kuvailevan kirjallisuuskatsauksen aiheeseen vahvistui, kun lähdeaineiston hakuprosessissa paljastui, että aihetta ei ole tutkittu tai vertailtu erilaisten urheilulajien näkökulmasta, vaan usein tutkimukset rajautuivat aina yhteen tiettyyn urheilulajiin. Erityisesti aikaisemmista tutkimuksista näkyi, että tutkimuksissa ei tutkittu urheilijan urheilullista puolta tai urheilijan kehitysmahdollisuuksia datan hyödyntämisessä oikeastaan lainkaan, vaan enemmänkin tutkimukset datan käytöstä urheilussa viittasivat kaikki urheilufanien, valmentajien, urheiluorganisaatioiden tai urheilutapahtumien datan hyödyntämiseen.

Koska datankerääminen ja data-analytiikan sekä teknologian hyödyntäminen huippu-urheilussa ei ole ollut aikaisempien tutkimusten nimenomaisena tutkimuskohteena, on perusteltua tarkastella teoriapohjaa ja käsitteistöä seuraavanlaisten teemojen kautta: big data, tekoäly, data-analytiikka, urheiluanalytiikka, IoT ja urheiluteknologialaitteet. Opinnäytetyön tarkoituksena on luoda eheä ja kokonaisvaltainen kuvaus edellä mainittujen teemojen suhteesta huippu-urheiluun ja siitä kerättävään dataan rajatuissa joukkueurheilulajeissa. Puutteellisten ja hajanaisten aikaisempien tutkimusten vuoksi teoriapohjaa ei ole mahdollista johtaa suoraan yhdestä diskurssista vaan on tarkasteltava niitä myöhemmin havainnollistamalla tavoilla. Kerään, yhdistelen ja analysoin aikaisempien tutkimusten perusteella oman kuvailevan kirjallisuuskatsauksen kannalta valideimmat tutkimustulokset.

Ymmärtääkseen tämän opinnäytetyön riittävällä tasolla, on välttämätöntä perehtyä keskeisiin käsitteisiin, jotka edellisessä kappaleessa mainitsin teemoina. Nämä kaikki käsitteet ovat avainasemassa, kun tarkastellaan data-analytiikan ja teknologian hyödyntämistä huippu-urheilussa.

3.1 Big Data

Vielä muutamia vuosia sitten emme voineet kuvitellakaan, että voisimme kerätä ja analysoida dataa yhtä laajoja määriä kuin tänä päivänä. Käytettävissä oleva tietomäärä nykypäivänä on niin massiivinen, että vuoteen 2020 ennustettiin, että joka sekunti luodaan

1,7 megatavua uutta tietoa (Bernard 2016, 1-2). Näin suurien datamassojen keräämistä ja analysointia kutsutaan termillä big data. Bernardin (2016, 2) mukaan kaikki tämä tieto ei synny ainoastaan sähköposteista, Facebookista, Whatsapp-viesteistä tai Twitteristä, vaan myös biljoonasta digitaalisesta kuvasta. Videot ja niistä saatava data tulee olemaan tässä tutkielmassa keskeisessä roolissa, sillä tänä päivänä urheilusta saatavaa dataa on tarjolla esimerkiksi digitaalisen kuvan muodossa.

Videoiden lisäksi urheilijoista kerätään dataa erilaisilla mittaavilla sensoreilla, kuten esimerkiksi sykemittareilla. Sensoreita on myös esimerkiksi älypuhelimissa, jotka kertovat missä olemme (GPS), kuinka nopeasti liikumme (kiihtyvyyssmittari), millainen sää ulkona on (barometri) ja mitä kohtaa kosketamme älypuhelimien näytöllä (kosketusanturi). (Bernard 2016, 2.) Big data käsitettä selventäessä on tärkeää ymmärtää, että älypuhelimet eivät ole ainoita teknologialaitteita, joihin ihmiset jättävät dataa itsestään. Kotitalouksista löytyy myös älytelevisiot, älykellot, älykkäät vedenkeittimet, jääkaapit, tennismailat ja jopa hehkulamput. Vuonna 2016 julkaistussa kirjassa Bernardin (2016, 4) mukaan meillä on vuoteen 2020 mennessä yli 50 miljardia laitetta, jotka ovat yhteydessä internetiin. Bernard (2016, 4) on vakuuttunut myös siitä, että toisin kuin mikään muu trendi, big data, vaikuttaa tällä hetkellä lähes kaikkeen, mitä me ihmiset teemme.

Tässä opinnäytetyössä tutkitaan, kuinka big dataa hyödynnetään huippu-urheilussa erilaisissa urheilulajeissa. Bernard (2016, 59) on tutkinut myös big dataa urheilun näkökulmasta ja hänen mukaansa big dataa voidaan hyödyntää urheilussa esimerkiksi urheilijoiden unen, palautumisen ja harjoittelun mittaamisessa. Big data on yleisenä teknologian käsitteenä helposti löydettävissä ja ymmärrettävissä, mutta sen vaikutusta urheilussa ei olla kuitenkaan tutkittu niin laajasti, että sitä pystyisi selittämään täysin urheilun näkökulmasta. Käsitteenä big data tulee esiin usein urheiluanalytiikan tutkimuksissa ja se mielletään näissä tutkimuksissa täysin yleismaailmalliseksi käsitteeksi, niin kuin kaikissa dataa käsittelevissä tutkimuksissa.

3.2 Data-analytiikka

Termistä data-analytiikka tuli suosittu 2000-luvun alussa. Data-analytiikalla tarkoitetaan tietokonejärjestelmien käyttöä ja hyödyntämistä suurten tietojoukkojen (big data) analysoinnin ja päätösten teon tueksi. Data-analytiikka on hyvin monitieteinen ala, joka on omaksunut näkökulmia monilta muilta tieteenaloilta, kuten tilastotiede, koneoppiminen, mallintaminen, järjestelmäteoriatiede ja tekoäly. (Runkler 2016, 26.)

Tyypillinen data-analyysiprojekti jaetaan useisiin eri vaiheisiin. Tiedot arvioidaan ja valitaan, puhdistetaan ja suodatetaan sekä visualisoidaan ja analysoidaan, jonka jälkeen analyysitulokset tulkitaan ja lopulta arvioidaan. (Runkler 2016, 26–27.)

Data-analytiikan avulla voidaan löytää asiaankuuluvia tietoja, rakenteita ja malleja siitä kuinka saada uutta tietoa, sekä tunnistaa syyt ja seuraukset. Data-analytiikan avulla voidaan myös ennustaa tulevaa kehitystä tai ehdottaa optimaalisia päätöksiä. Jotta data-analytiikasta saadaan mahdollisimman relevanttia tietoa, tarvitaan avuksi malleja ja algoritmeja tietojen keräämiseen, esikäsittelyyn, analysointiin ja arviointiin eri aloilta, kuten tilastoista, koneoppimisesta, mallientunnistuksesta, järjestelmäteorioista, operaatiotutkimuksista tai tekoälystä. (Runkler 2016, 8.)

Urheilun näkökulmasta data-analytiikasta puhuessa voi törmätä usein myös termiin urheiluanalytiikka, joka ei kuitenkaan ole täysin sama asia kuin data-analytiikka. Data-analytiikan avulla urheilussa pyritään tunnistamaan syyt ja seuraukset. Syiden ja seurauksien tunnistaminen urheilussa on mahdollista siitä kerättyä dataa hyödyntämällä ja analysoimalla. Ennen mekaanisesti kerätyt tilastot ottelutapahtuman aikana olivat ainoa data, mitä urheilusta saatiin. Teknologian kehittyminen on mahdollistanut, että tilastoja ei tarvitse enää kerätä mekaanisesti, vaan teknologialaitteet tekevät sen ihmisten puolesta. Data-analytiikkaa on tänä päivänä helpompi hyödyntää urheilussa kuin koskaan ennen.

3.3 Urheiluanalytiikka

Termi urheiluanalytiikka (Sports analytics) tunnetaan myös yksinkertaisesti urheilutilastointina (Statistics in sports). Urheiluanalytiikka on datan keräämistä, listaamista ja hallitsemista. Urheilujoukkueet, pelaajat ja valmentajat pyrkivät urheiluanalytiikan avulla ennakoimaan ja löytämään arvokasta tietoa heidän toimintansa ja päätöksentekonsa parantamiseksi. Datan keräämisen teknologioita ja mahdollisuuksia on tänä päivänä valtavasti, mutta urheilussa datan keräämisellä ei ole merkitystä, jos analytiikan menetelmät eivät ole selviä. Urheiluanalytiikka voidaan myös mieltää tieteelliseksi alaksi, joka käsittelee kokoelmaa aiemmasta ja nykyisestä urheiludatasta sekä sen keräämisestä ja analysoimisesta. Tällainen urheiluanalytiikka voi auttaa urheiluseuroja ennustamaan tulevia tilanteita etukäteen tai tekemään oikeita päätöksiä kauden aikana. Urheiluanalytiikassa ja datan mittaamisessa urheilussa on myös ongelmansa. Yleisesti urheiludata voi olla epäsäännöllistä ja hajanaista, koska pelaajien urat eivät ole aina monivuotuisia, eivätkä pelaajat pysy samassa joukkueessa montaa vuotta. Tämä tekee urheilusta, joukkueista ja pelaajista kerätyn datan analysoimisesta -eli urheiluanalytiikasta- erittäin haastavaa (Sarlis & Tjortjis 2020, 1-2).

3.4 Liikuntateknologialaitteet

Liikuntateknologialaitteita (Sports Technologies) on maailmassa valtavasti, mutta yleisimpiä ja tunnetuimpia ovat puettavat sensorijärjestelmät, kuten sykemittarit ja

älykellot. Lemmettely (2016, 1) kirjoittaa tutkimuksessaan, että puettavat sensorijärjestelmät ovat järjestelmiä, joihin tuotetaan tietoa sensoreiden avulla ja tallennettu tieto käytetään hyväksi, eli toisin sanoen analysoidaan, niille tarkoitetuissa sovelluksissa.

Liikuntateknologianlaitteet ja urheilusuoritusten mittaamisen yleistymisen on tapahtunut matkapuhelinten ja älykellojen yleistymisen vuoksi. Sensorijärjestelmien ja teknologialaitteiden kehittyminen on tarjonnut mahdollisuuden monimutkaisempienkin ongelmien mittaamiseen, kuten liikkeen mallintamiseen, voimatasojen mittaamiseen, vammojen ehkäisyyn ja jopa pelitaktiikan analysoimiseen. (Lemmettely 2016, 3.)

Suoritusten mittaaminen on mahdollista myös urheiluvälineiden avulla. Sensoreita laitetaan esimerkiksi jääkiekkomailan varren tai tennismailan sisään, jolloin on mahdollista mitata esimerkiksi mailan taipumaa, alakäden sijaintia ja mailan liikettä. Lopuksi liikkeet analysoidaan ja luokitellaan, jonka jälkeen urheilijalle voidaan esittää tarvittavat muutokset harjoitteluun. Tenniksessä on käytetty paljon videoanalyysiä apuna, jolloin voidaan mitata ja tarkkailla esimerkiksi kahden tai jopa kolmen eri muuttujan vaikutusta toisiinsa. (Lemmettely 2016, 10.)

Liikuntasensoreiden kehitystä ovat lisänneet tiedonsiirtoverkkojen laajentuminen ja pilvipalveluiden yleistymisen. Näillä keinoilla tietojen analysoiminen ja jakaminen on entistä helpompaa ja nopeampaa. Urheilusuoritusten ja urheilijoiden mittaaminen ei ole mikään yksinkertainen tehtävä. Urheilusuoritusten liikkeet, iskut, vaihtuvat lämpötilat ja kosteus vaihtelut aiheuttavat sensoreille haasteita. Nämä haasteet on otettava huomioon virheiden minimoimiseksi ja kestävien laitteiden kehittämiseksi. (Lemmettely 2016, 2.)

Liikuntateknologialaitteilla ei tarkoiteta pelkästään sykemittareita tai sensoreita vaan kaikkia teknologialaitteita, joilla on mahdollista seurata urheilua tai urheilijaa saaden siitä jonkinlaista dataa. Videokuvasta onkin kehittynyt yksi tärkeimmistä datankeräämisen lähteistä huippu-urheilussa koneoppimisen ja tekoälyn myötä. Yksi mielenkiintoisimmista teknologialaitteista urheilussa on SportsVU -kamerat, jotka kehitettiin vuonna 2005 Israelissa optiseksi tunnistuslaitteeksi, joilla voitiin tunnistaa ohjuksia (Cai & Zhao 2021). Nyt SportsVU kameroita käytetään laajasti urheilussa ja varsinkin joukkuelajeissa. NBA:ssa SportsVU kamerat otettiin käyttöön vuonna 2009 ja nyt niillä voidaan seurata pelaajien jokaista liikettä, kuten syöttöä, laukausta, pallon kosketusta ja kaikkea sitä, millä on vaikutusta itse peliin (Cai & Zhao 2021).

3.5 Tekoäly

Tekoälyä (Artificial Intelligence) kutsutaan usein myös AI:ksi ja sillä tarkoitetaan yleismaailmallisesti koneen kykyä jäljitellä inhimillistä ajattelua (Randén 2017). ”Tekoälyksi määritellään yleensä ratkaisu, joka oppii ja kykenee ratkomaan ongelmia, jotka eivät ole ratkaistavissa rutiininomaisella laskennalla,” kirjoittaa Randén (2017) blogitekstissään. Randén esittää myös, että tekoäly voidaan jakaa kolmeen kategoriaan: tekoäly, yleinen tekoäly (Artificial General Intelligence), joka tunnetaan myös termillä vahva tekoäly (Strong AI) ja sovellettu tekoäly (Applied AI), joka tunnetaan myös termillä heikko tekoäly. Vahvalla tekoälyllä tarkoitetaan koneen kykyä jäljitellä kokonaisvaltaisesti inhimillistä älykkyyttä ja koneen kykyä kehittää tietoisuutta. Sovelletulla tekoälyllä eli heikolla tekoälyllä tarkoitetaan, että sillä pystytään tuottamaan ratkaisuja joihinkin tarkasti rajattuihin ongelmiin. Vuonna 2017 tekoälysovellukset olivat vielä kaukana vahvasta tekoälystä. (Randén 2017.)

Randén (2017) kuvailee, että koneoppiminen (Machine Learning, ML) on yksi keskeisistä käsitteistä tekoälyssä ja termiä käytetäänkin joskus synonyyminä tekoälylle.

Koneoppiminen on tietojenkäsittelyn ala, jossa tietokone ohjelmoidaan oppimaan ja tunnistamaan erilaisia malleja ilman, että niitä erikseen syötetään tietokoneelle.

Tekoälysovellusten kehittämisen haasteellisin vaihe on saada tietokone tunnistamaan erilaisia malleja. (Randén 2017.) Koneoppimisen avulla pyritään löytämään isosta määrästä dataa säännönmukaisuudet, joita ihminen ei voisi löytää. Koneoppimisessa on kyse siitä, kuinka rakentaa tietokonejärjestelmä, joka paranee automaattisesti kokemuksen kautta. Järjestelmien ja koneiden koneoppimista on mahdollista kehittää algoritmeilla, suurilla datamäärillä tai tehokkailta tietokoneympäristöillä. (Pan, Vasilakos, Wang & Zhou 2017, 1.)

Urheilun näkökulmasta koneoppimisella on valtavat kehitysmahdollisuudet sieltä saatavien suurien datamassojen takia. Mitä enemmän järjestelmät ja koneet saavat tietoa opittavaksi, sen parempi mahdollisuus niillä on huomata säännönmukaisuuksia. Pan, Vasilakos, Wang ja Zhou (2017, 1) mainitsevat artikkelissaan koneoppimisen mullistuneen viime vuosina osittain suurten datamäärien myötä. Isojen datamäärien myötä koneoppimisesta on tullut paljastavampaa, tarkempaa ja nopeampaa kuin koskaan ennen (Pan, Vasilakos, Wang & Zhou 2017, 1). Urheilussa tekoälyä ja koneoppimista on hyödynnetty viime vuosina myös ennustaviin ja simuloiviin ratkaisuihin. Amazon Web Services pystyy rakentamaan tuhansia erilaisia ennustavia malleja pelitapahtumista. Ennustavilla koneoppimisen ja tekoälyn malleilla tehdään skenaarioita pelitapahtumista, ja niillä voidaan ennustaa esimerkiksi millainen pelitapa tietyssä pelitilanteessa toimisi parhaiten. (Ghasem, Saxon & Valenzuela 2021, 438.)

Aineistonhakuprosessissa oli löydettävissä tutkimuksia, joissa tutkittiin koneoppimisen ennustavia malleja, ja niissä pyrittiin ennustamaan pelien lopputuloksia ja vedonlyöntikertoimien muuttujia. Tämä opinnäytetyön aineisto on rajattu urheilun näkökulmaan, eikä tämä käsittele tekoälyä tai koneoppimisen ennustavia malleja lopputuloksen kannalta, vaan urheilun kannalta. Toisin sanoen ei lopputuloksen tietäminen auta pelaajaa tai joukkuetta voittamaan tai pelaamaan paremmin, mutta jos koneoppimisen avulla voidaan kertoa pelaajille, että kiekkoa ei tulisi laukoa paikasta X vaan todennäköisempää on, että se menee maaliin, jos kiekko ammutaan paikasta Y; voi se kehittää pelaajaa paremmaksi pelaajaksi ja on mahdollista, että se korreloi myös lopputuloksen kanssa.

3.6 IoT ja IoT-laitteet

IoT-termi (Internet of Things) on keksitty vuonna 1999. Termin keksi Kevin Ashton ja sillä tarkoitettiin radiotaajuisten etätunnistuksen (RFID) käyttöä tiedon etäluvussa ja tallentamisessa. Myöhemmin IoT on saanut erilaisia määritelmiä, jotka ovat olleet sidottuna verkon (web) kehittymiseen. (Mäntynen 2018, 5.)

IoT toimii yhdenlaisena datan lähteenä. Datasta pyritään tuottamaan laadukasta tietosisältöä ja hyödyntämään sitä yritysten liiketoiminnassa. (Mäntynen 2018, 5.) Myös urheiluseurat ovat alkaneet nähdä hyötyä IoT:n yhdistämisestä data-analytiikkaan ja pelaajien seurantalaitteisiin.

Videotallenteiden ja perinteisen manuaalisen tilastoinnin lisäksi monessa urheilulajissa hyödynnetään IoT-laitteita. Urheilussa IoT-laitteilla tarkoitetaan dataa kerääviä laitteita, jotka ovat yhdessä yhdistettynä internettiin muodostaen näin kokonaisuuden, jota kutsutaan laitteiden verkostoksi. (Ryhänen 2020, 9–10.) IoT-laitteita, joita urheilussa hyödynnetään ovat esimerkiksi erilaiset rannekkeet, rintahihnat, laastarit ja älykellot. Laitteilla voidaan seurata sykettä, kehon massaa, aktiivisuustasoa ja kuormitusta. Näitä laitteita voidaan käyttää myös harjoittelun aikana ja elämäntapojen, kuten unen tai levon mittaamisessa. (Miller 2015.)

IoT-laitteiden avulla urheiluseurat pystyvät keräämään laadukasta dataa pelaajista ja pelitapahtumista. IoT-laitteet ovat siis yhteydessä yhdistettynä verkkoon, mutta ovat myös yhteydessä toisiinsa erilaisten sensoreiden tai sirujen avulla. Sensoreita on laitettu moniin älykkäisiin laitteisiin, kuten esimerkiksi älypuhelimiin tai sykemittareihin ja sensorit mahdollistavat internetyhteyden avulla sen, että näistä laitteista tulee älylaitteita. (Mäntynen 2018, 18.) Urheilussa sensoreiden lisäksi käytetään siruja, jotka kiinnitetään pelaajien suojuksiin ja pelivälineeseen, kuten kiekkoon. Sirujen avulla voidaan seurata

reaaliaikaisesti pelaajien sekä pelivälineen liikettä kentällä. Siruteknologiasta esimerkkinä on myöhemmin tutkielman tuloksissa esitelty Suomen jääkiekon pääsarjassa Liigassa käytettävä älykiekko. (Wisehockey Oy 2020.)

IoT-laitteet pystyvät hyödyntämään analytiikkaa sekä toimintatietoja toimiakseen nopeammin kuin yksikään ihminen. IoT-laitteiden reaktiot ja toiminnot tapahtuvat täysin itsestään ilman ihmisten tietoisuutta asiasta. (Mäntynen 2018, 5.) Urheilun näkökulmasta IoT-laitteiden ja teknologian kehittyminen on mahdollistanut suurten datamäärien keräämisen nopea tempoisista peleistä ja nopeasti liikkuvista pelaajista. Tänä päivänä IoT-laitteita hyödynnetään datankeräämisessä urheilussa juuri niiden nopeuden ja automaattisen toiminnan vuoksi.

4 Data-analytiikan ja teknologian hyödyntäminen huippu-urheilussa

Dataa voidaan hyödyntää urheilussa esimerkiksi otteluiden analysointiin, pelistrategian- ja harjoittelun suunnitteluun, loukkaantumisten ehkäisyyn, suorituskyvyn hallintaan ja ennustamiseen, sekä urheilujoukkueiden rakentamiseen. Datankeräämisen lisäksi tiedontarpeiden tunnistaminen huippu-urheilussa on olennaista, koska dataa on mahdollista kerätä niin valtavia määriä. (Ryhänen 2020, 8.) Urheilusta kerättävä data voi olla määrällistä tai laadullista ja sitä on saatavilla monissa eri muodoissa, kuten numeraalisena, videotallenteena tai kirjallisena. Videotallenteiden ja perinteisen manuaalisen tilastoinnin lisäksi lähes kaikissa joukkueurheilulajeissa hyödynnetään tänä päivänä IoT-laitteita, joiden avulla pelaajista ja ottelutapahtumista on mahdollista kerätä dataa. (Miller 2015.)

Tiedontarpeet voidaan jakaa viiteen luokkaan: ensimmäisessä urheiludataa ei hyödynnetä oikeastaan lainkaan, toisessa tehdään ennustavia päätöksiä tilannekohtaisesti inhimillisen ajattelun perusteella, kolmannessa asiantuntijat tekevät ennustavia päätöksiä perustuen aikaisempaan näyttöön ja tietoon, neljännessä hyödynnetään laajasti tilastoja päätöksenteonprosesseissa ja viidennessä hyödynnetään tiedon louhintaa (Data mining) päätöksenteonprosessien apuna. Luokat kolme, neljä ja viisi ovat luokkia, joissa hyödynnetään data-analytiikkaa. Luokat yksi ja kaksi perustuvat suoritusten ja tilanteiden tarkasteluun, jossa dataa ja teknologiaa ei hyödynnetä riittävällä tasolla ollakseen laadukasta tämän päivän huippu-urheilussa pärjäämiselle. (Ryhänen 2020, 8.)

4.1 Suorituskyvyn mittaaminen joukkueurheilussa KPI-mittareiden avulla

Yksi eniten käytetyistä datankeräysmenetelmistä ja urheiluanalytiikan keinoista urheilujoukkueen suorituskyvyn mittaamiseksi ovat KPI-mittarit, eli Key Performance Indicators. KPI tarkoittaa tärkeimpiä suorituskykyyn vaikuttavia tekijöitä, joita on urheilussa paljon ja joiden mittaamiseen tarvitaan monia erilaisia työkaluja. (Ryhänen 2020, 9.) KPI-mittareita käytetään etenkin joukkueurheilussa ja niillä mitataan esimerkiksi pelaajien välistä etäisyyttä kentällä eri tilanteissa tai eri pelipaikkoja pelaavien pelaajien suorituskyvyvaatimuksia. KPI-mittareilla on mahdollista mitata myös pelikenttien eri osa-alueiden tapahtumia tai eri osa-alueiden vaatimustasoja, sekä pelaajien välisiä optimaalisia kulmia erilaisissa haastavissa pelitilanteissa. Pelaajista on mahdollista mitata KPI-mittareiden avulla esimerkiksi aikaa kauan he ovat olleet kentällä vapaana pelattavissa, jotta heille olisi ollut mahdollista syöttää peliväline, pelivälineen hallussapitoaikaa, pelaajien nopeuksia kentän eri osa-alueilla ja tilanteissa sekä pelaajien fyysisiä ominaisuuksia, kuten esimerkiksi syke- tai laktaattitasoja. (Ryhänen 2020, 9.)

NBA:ssa (National Basketball Association) on käytetty KPI-mittareita enemmän joukkuepelejä parantavan datankeräämiseen, kuin pelkästään pelaajien yksilöllisen datankeräämiseen. NBA:ssa KPI-mittareilla tarkastellaan seuraavia asioita: kuinka tehokas joukkue on tekemään pisteitä, eli koreja verrattain vastustajaan (Effective Field Goal Percentage, eFG%), kuinka monta kertaa pallo menetetään niin, että vastustaja saa vastahyökkäyksen tai päinvastoin, kuinka monta kertaa oma joukkue riistää pallon niin, että oma joukkue pääsee vastahyökkäykseen (Turnover Percentage, TOV%), kuinka monta levypalloa hävitään ja voitetaan (Offensive and Defensive Rebound Percentage, ORB%, DRB%), sekä vapaaheittoerointa, eli kuinka usein oma joukkue saa vapaaheiton ja kuinka usein se menee koriin (Free-throw factor, FT / FTG). (Dehesa ym. 2019, 92–93.)

KPI-mittareiden avulla organisaation toimintaa voidaan kohdentaa siihen mikä on tärkeää ja sen avulla voidaan parantaa kokonaisvaltaista ymmärrystä, päätöksentekoa ja toteutusta. Tasaisin väliajoin ja oikein tehty mittaus sekä mittauksien analysointi auttavat organisaatiota valmistautumaan tulevaisuuteen, mikä vähentää virheitä. (Parmenter 2015, 18.) Ennen kaikkea urheilujoukkueen näkökulmasta pelaajien ja joukkueen motivoiminen nousee suureen rooliin datankeräämisen hyödyissä ja se, että pelaajilla ja valmentajilla on yhteinen ymmärrys, mitkä teot kentällä vaikuttavat itse tärkeimpään eli ottelun voittamiseen.

Alam ja Almulla (2020, 2) kirjoittavat, että viime vuosina jalkapallo-otteluihin on tullut monia uusia moderneita teknologian tekniikoita, jotka ovat parantaneet pelin laatua ja auttanut pelaajien suoritusten seurantaan otteluiden aikana. Tällaisia uudistuksia on esimerkiksi erilaiset pelaajien käyttämät seurantalaitteet, monikamera tekniikka (multi-camera) ja videotuomaritukijärjestelmä (VAR). Tällaiset tekniikat mahdollistavat sen, että otteluista saadaan kerättyä isoja määriä dataa, jota voidaan myöhemmin käyttää pelaajien ja joukkueen suorituskyvyn ymmärtämiseen ja ottelutuloksen analysoimiseen dataan perustuvalla tavalla. (Alam & Almulla 2020, 2.) Teknologian, tekoälyn ja datankeräysmenetelmien kehittyminen on tehnyt KPI-mittareiden käyttämisestä entistä helpompaa. Enää otteluista ja pelaajista ei tarvitse kerätä tietoja mekaanisesti.

4.2 Data-analytiikan ja teknologian hyödyntäminen Liigassa

Puettavien IoT-laitteiden lisäksi teknologiaa hyödynnetään datankeräämisessä urheilukentille asennettavien sensoreiden avulla sekä pelivälineisiin asennetuilla siruilla. Suomessa jääkiekon pääsarjassa Liigassa sirumenetelmää on hyödynnetty kiekkoon sekä pelaajiin asennetuilla bluetooth-siruilla, mitkä ovat yhteydessä jäähalliin asennettuihin lukuisiin sensoreihin. Näiden sirujen ja halliin asennettujen sensoreiden avulla saadaan reaaliaikaista dataa analysoitavaksi pelitilanteista. (Ryhänen 2020, 10.) Kiekkoa ja sen sisään asennettua sirua kutsutaan älykiekoksi. Älykiekko teknologia perustuu

suomalaisen Quuppa-yhtiön RTLS-ratkaisuun (Real-Time Location System) sekä Suomessa valmistettuihin paikannuslaitteisiin. (Wisehockey Oy 2020.) Wisehockey Oy:n toimitusjohtaja Tomi Mikkonen kertoo Wisehockeyn (2020) julkaisemassa artikkelissa, että älykiekon lisäksi kaikkien joukkueiden pelaajilla on bluetooth-sirut ommeltuna pelipaitoihin. Wisehockeyn järjestelmän sensorit ja palvelimet lukevat sirun älykiekosta sekä pelaajista, jonka jälkeen analyysi sirujen datasta tapahtuu järjestelmässä. Tämän data-analytiikkajärjestelmän avulla kentällä olevat pelaajat sekä älykiekko muuttuvat datapisteiksi ja lopulta pilvijärjestelmän avulla kaikesta kentällä tapahtuvasta tulee tilastoita, graafikoita, kuvia ja videoita. (Wisehockey Oy 2020.) Kaikki edellä mainitut datapisteet kertovat lopulta meille enemmän analytiikkaa, kuin mekaanisesti ihminen pystyisi koskaan ottelun tapahtumista erottamaan.



Figure 3: The smart puck looks and weighs the same as an ordinary puck

Kuva 1. Wisehockey Oy:n Älykiekko ja kiekkoon laitettava siru (Wisehockey Oy 2020)



Figure 2: Player tags

Kuva 2. Wisehockey Oy:n kehittämät pelaajien paitoihin ommeltavat sirut (Wisehockey Oy 2020)

Wisehockey Oy:n järjestelmä Liigaan on kattava ja se tallentaa dataa pelaajista, tuomareista ja itse pelistä. Wisehockeyyn järjestelmä kerää ja analysoi dataa pelin aikana pelaajien nopeuksista, matkasta, jonka pelaajat luistelevat pelin aikana, yksittäisen pelaaja kokonaispeliajasta ja pelaajan vaihtojen kestoista. Järjestelmällä mitataan myös, kuinka kauan kullakin yksittäisellä pelaajalla on kiekko hallussaan tai kummalla joukkueista kiekko on enemmän hallussa pelin aikana. Järjestelmällä kerätään dataa myös siitä, kuinka hyvin pelaajien syötöt onnistuvat pelin aikana ja kuinka kovaa pelaajat laukovat tai syöttävät kiekkoa, kumpi joukkue voittaa enemmän aloituksia ja kuinka monta laukausta joukkue ampuu pelin aikana ja osuvatko laukaukset maalia kohti vai menevätkö ne ohitse. Wisehockeyyn järjestelmä antaa myös reaaliaikaista tietoa pelinaikana rangaistuksista, ylivoimapelistä, pelaajien plus-miinus-tilastoista ja jopa siitä kumpi joukkue hallitsee peliä. (Wisehockey Oy 2020.)

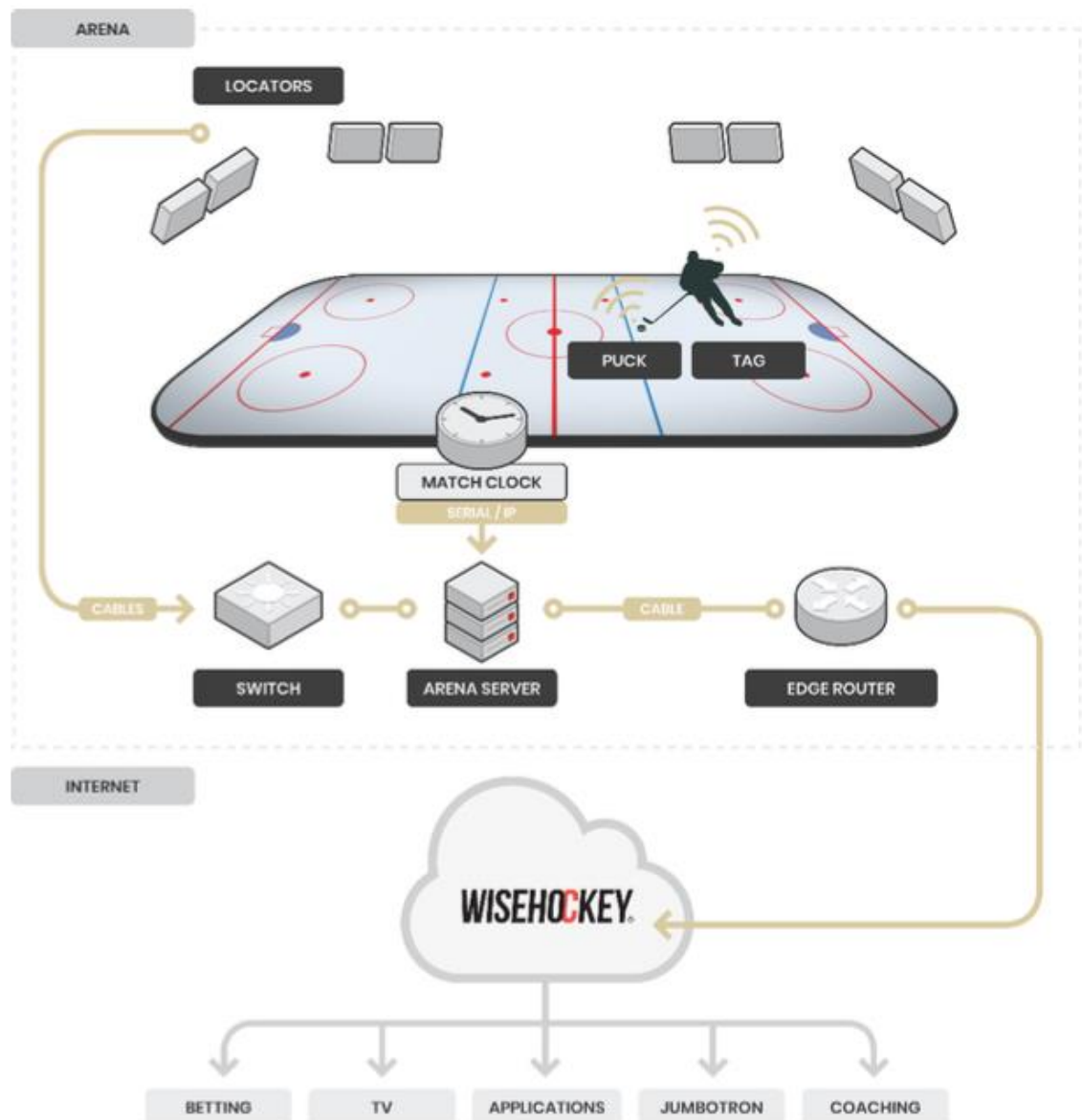


Figure 1: Hardware infrastructure

Kuva 3. Wisehockey Oy:n kehittämä sensorijärjestelmä jäähalliin asennettuna (Wisehockey Oy 2020)

4.3 Data-analytiikan ja teknologian hyödyntäminen NHL:ssä

NHL (National Hockey League) tekee yhteistyötä SAP:n kanssa ja hyödyntää heidän datankeräys järjestelmiään. NHL kirjoitti vuonna 2015 pitkän sopimuksen SAP:n kanssa ja siitä lähtien NHL on alkanut hyödyntämään IoTsports-teknikkaa datan keräämisessä ja datan analysoimisessa. SAP:n IoTsports-teknikan avulla NHL on alkanut julkaista paranneltuja tilastoja esimerkiksi pelaajien laukausyriksistä, maalia kohti menevästä laukausprosentista, maalivahtien torjuntaprosenteista ja kolmestakymmenestä muusta laajennetusta tilastosta. (Mincev 2021, 35.)

NHL kauden 2020–2021 palkkakatto oli asetettu 81,5 miljoonaan dollariin. Tämä tarkoittaa siis sitä, että jokaisella 31 joukkueella NHL:ssä on käytettävissään kyseinen summa heillä pelaaviin pelaajiin, eli pelaajabudjettiin. (Mincev 2021, 33.) Toisin kuin useimmissa muissa urheilusarjoissa, NHL:ssä on käytössä edellä mainittu palkkakatto. Palkkakatto on kehitetty minimoimaan tasoeroja rikkaiden ja köyhien joukkueiden välillä. Esimerkiksi New York Rangers on NHL:n rikkain joukkue ja sen arvo on 1650 miljoonaa dollaria, kun taas verrattain Arizona Coyotesin arvo on vain 300 miljoonaa dollaria. (Mincev 2021, 34.) Yksi NHL:ssä todettu ongelmakohta on, että pelaajille on maksettu liikaa palkkaa heidän oikeaan arvoonsa nähden. Monet pelaajat ovatkin joko yliarvostettuja tai aliarvostettuja NHL:ssä, eli toisin sanoen ylipalkattuja tai alipalkattuja. (Mincev 2021, 35.) Edellä mainittuja tilanteita välttääkseen NHL-joukkueet ovat hyödyntäneet dataa myös pelaajien palkkojen arvioimisessa. Yksi käytetyin menetelmä on pelaajien palkkojen vertaaminen heidän plus-miinus-tilastoonsa (Mincev 2021, 41) eli tilastoon, joka määräytyy sen mukaan, kuinka useasti pelaaja on ollut kentällä, kun oma joukkue on tehnyt maalin verrattuna siihen, kuinka usein pelaaja on ollut kentällä, kun vastustaja on tehnyt maalin. Plus-miinus-tilastossa huomioidaan vain tasakentällisin tehdyt maalit, poikkeuksena kuitenkin on alivoimalla tehty maali, joka myös antaa plusmerkinnän pelaajan plus-miinus-tilastoon (NHL 2021.) Plus-miinus-tilastoa tulkitaan yksinkertaisuudessaan niin, että mitä suurempi arvo pelaajalla tässä tilastossa on, sitä suurempi pelaajan panos on ollut. Plus-miinus-tilaston lisäksi palkkaa voidaan suhteuttaa myös esimerkiksi pelaajan tekemiin pisteisiin, syöttöpisteisiin, tehtyihin maaleihin, päästettyihin maaleihin tai pelaajan ikään (Mincev 2021, 41). NHL pelaajien ikää on tutkittu myös laajemmin, kuin vain palkkojen muodossa. Dataa on tutkittu ja analysoitu niin, että hyökkääjät sekä puolustajat ovat parhaimmillaan 26-vuotiaina. Lisäksi olemassa olevasta datasta on analysoitu, että hyökkääjien suorituskyyky on lähellä huippua kymmenen vuoden ajan ja puolustajilla kahdentoista vuoden ajan. (Swartz 2017, 12.)

NHL:n alkuaikoina, ja tälläkin hetkellä alemmissa jääkiekkosarjoissa, käytetään apuna tilastointia. Tilastointia pelistä suorittavat tilastohenkilöstö ja kaikki tilastointi tapahtuu mekaanisesti. Tietoja pelaajista ja pelistä syötetään paperille tai tietokoneelle, kun tarkasteltavana on pelaajien laukaukset, torjunnat, maalit, hyökkäysalueelle sisäänviennit ja aloitukset. Kun peli modernisoitui ja pelejä alettiin kuvata videolle enemmän, tilastointia alettiin tekemään jälkikäteen videoilta. Jälkikäteen videolta katsottuna tilastointi tietoja saatiin huomattavasti enemmän. Tänä päivänä NHL:ssä kotijoukkueella on velvollisuus kerätä reaaliaikaista dataa peleistä ja ladata data eteenpäin NHL.com -sivustolle, josta kuka tahansa voi seurata pelin tilastoja reaaliajassa. Useat NHL-joukkueet ovatkin investoineet omaan analytiikkahenkilöstöön. (Swartz 2017, 3-4.) Kaikkeaa edellä mainittua tilastointia ja dataa, mitä NHL.com -sivustolle lisätään, kutsutaan RTSS-dataksi, eli Real-Time Scoring System (Curro & Shuckers 2013, 1). Pelattavat pelitiedot haetaan

automaattisesti järjestelmän avulla RTSS-tietokannasta ja tiedot tallentuvat tiedostoihin, joita käsitellään R-ohjelmointikielen avulla. Tyypillisesti yhdestä ottelusta syntyy 400 reaaliaikaista tapahtumaa ottelua kohden, mikä vastaa yhtä tapahtumaa yhdeksän sekunnin välein. (Swartz 2017, 5.)

Vaikka NHL investoi analytiikkaan 2010-luvun alussa, on se kuitenkin aloittanut SportVU-järjestelmän käytön paljon NBA:ta myöhemmin. NHL ei ollut vielä kaudella 2016–2017 ottanut SportVU-järjestelmää käyttöönsä, koska se ei ollut mahdollista kaikille sarjassa pelaaville joukkueille. (Swartz 2017, 5.) Jos osa NHL-joukkueista olisivat saaneet käyttöönsä SportVU-järjestelmän, olisivat kyseiset joukkueet saaneet tästä huomattavan edun. Kun SportVU-järjestelmät eivät päässeet osaksi NHL:n pelikenttiä tulivat kilpailevat tekniikat käyttöön. Sportvision kehitti sirut pelaajien paitoihin sekä kiekkoon. SportLOGiQ-yhtiö asennutti kunkin joukkueen areenalla yhden kameran, joka perustui koneoppimiseen ja optiseen tunnistusjärjestelmään. Järjestelmän nimi oli SPORTiQ. SPORTiQ:n käytössä huomattiin paljon yksityiskohtia pelin tapahtumista ja tarkkuus yksittäisestä tapahtumasta oli hyvä. Keskimäärin SPORTiQ-kamerat tallensivat tapahtumia 1,2 sekunnin välein kentältä. (Swartz 2017, 5.) SPORTiQ-kamerat tallensivat myös x- ja y-koordinaatit, jokaiselta jäällä olevalta pelaajalta, joka on tärkeä pelaajaseurantadatan osa-alue.

Pelaajien paikannusjärjestelmät ja puettavat teknologialaitteet tulivat NHL:ään vasta tammikuussa vuonna 2019. Uuden datankeräysmenetelmien ja teknologian avulla otteluista voidaan kerätä yli 25 kertainen määrä dataa. NHL kuitenkin sallii kerätyn datan käyttämisen ainoastaan pelistrategiassa ja pelaajien kehittämisessä. Edellä mainituilla laitteilla kerätyn datan hyödyntäminen pelaajan kanssa käytävissä sopimusneuvotteluissa on kiellettyä. (Räikkönen 2021, 19.)

Swartz (2017, 10) kirjoittaa tutkimuksessaan datan hyödyistä ja käsittelee aihetta viitaten jääkiekkjoukkueen taktiseen pelaamiseen, eli taktiikkaan. Taktiikka on päätöksiä, jotka ovat valmentajan hallinnassa, ja koska jääkiekko on peli, jossa on vaikea hallita tilanteita ja määrätä tapahtumia, taktiikkaa ei usein käytetä. Taktiikkaa ei olla käsitelty analyttisessä mielessä niin huolellisesti, kuin olisi ollut mahdollista ja Swartz pitääkin sitä avoimena ongelmana (Swartz 2017, 10). Yksi esimerkki taktisesta urheiluanalytiikasta jääkiekossa ja eritoten NHL:ssä, jota on tutkittu, on maalivahdin ottaminen pois kentältä joukkueen ollessa yhden maalin häviöllä, joka näin ollen mahdollistaa kuudennen kenttäpelaajan tulemisen kentälle. Tilanne on tuottanut kompromissin, jossa häviöllä oleva joukkue tekee todennäköisemmin maalin, mutta samalla johdolla oleva joukkue tekee todennäköisemmin maalin tyhjään maaliin. (Swartz 2017, 10.) Tämä kompromissi todistaa, että data ja urheiluanalytiikka eivät aina ole täydellistä ja oikeassa. Swartz (2017, 10) kirjoittaa, että maalivahdin pois ottamista maaliiltaan on tutkittu onnistuneesti ja yhdellä

maalilla häviöllä olevan joukkueen tulisi tehdä se kolme minuuttia ennen peliajan loppumista. Ennen tätä analyysiä NHL-valmentajat ottivat maalivahdin pois maaliltaan alle yhden minuutin kohdalla, mutta on havaittu, että analyysin myötä suuntaus maalivahdin maaliltaan pois ottamiseen on aikaistunut (Swartz 2017, 10).

Jääkiekko on analytiikalle vaikea laji, kuten jalkapallo ja koripallokkin. Ne ovat jatkuvia pelejä, joissa on paljon liikkuvia osia. Jääkiekon analytiikassa on vaikeaa ymmärtää sitä, mitä tapahtuu silloin, kun kiekko ei ole pelaajilla hallussaan, eli kiekollisen pelin ulkopuolella, koska sellaisia tilanteita ei yleensä analysoida. Näillä vaikeasti havaittavilla tapahtumilla voi olla syvälinenkin vaikutus pelin tuloksiin. Pelivälineen hallussapidon vaikeus lisää analysoimisen vaikeutta jääkiekossa. Jalkapalloon ja koripalloon verrattuna kiekon hallussa pitäminen pelinaikana on huomattavasti vaikeampaa. Tämä vaikeuttaa ennen kaikkea hyökkäystaktiikoiden suunnittelua ja toteutuksia. (Swartz 2017, 13.)

4.4 Data-analytiikan ja teknologian hyödyntäminen jalkapallossa

Jalkapallossa, erityisesti eurooppalaisissa sarjoissa, käytetään monia datankeräysmenetelmiä, joilla saadaan dataa pelaajien suorituksista pelin aikana. Yksi datankeräysmenetelmä on videokuva. Jalkapalloareenoilla on kameroita jopa 24:sta aina 70:een, jotka kuvaavat peliä mahdollisimman monesta eri kulmasta. (Alam & Almulla 2020.)

Jalkapallossa johtavia videodatan keräysjärjestelmiä ovat Stats Perform ja Opta. Nämä edellä mainitut järjestelmät keräävät pelaajista dataa otteluiden aikana monikamerajärjestelmällä. Monikamerajärjestelmä on tehokas ja tarkka tapa kerätä dataa pelaajien liikkeistä ja peliajasta kentällä. Jalkapallossa käytetään myös muita datankeräysmenetelmiä, kuten puettavia sensoreita pelaajilla ja pelivälineeseen kiinnitettäviä seurantalaitteita. (Alam & Almulla 2020.)

IoT-laitteiden kehityksen myötä urheiluun ovat tulleet myös ulkoiset yritykset, joita joukkueet voivat halutessaan palkata keräämään dataa ja analysoimaan sitä omilla järjestelmillään. Yksi tällaisista yrityksistä on jo aikaisemminkin mainittu yritys nimeltä SAP ja heidän järjestelmänsä nimeltä HANA. HANA on reaaliaikainen datankeräys- ja data-analytiikkajärjestelmä, jota esimerkiksi Saksan jalkapallomaajoukkue hyödynsi voittaessaan kultaa vuonna 2014 jalkapallon MM-kisoissa. SAP:n HANA-järjestelmää hyödynnettiin suoraan sovelluksessa nimeltä Match Insights, joka tarjosi Saksan valmentajille näkemyksiä heidän pelaajiensa suorituksista, mutta myös kaikista vastustajista. Sovellukseen kirjattiin muun muassa yksilöllisiä syöttöjä, potkuja, pelaajien nopeuksia ja liikkeitä kentällä. Dataa sovellukseen kerättiin reaaliajassa esimerkiksi on-sight kameroilla ja kamerasensoreilla. Todellinen kerätyn datan muutos nähtiin SAP:n

HANA-sovelluksen avulla, kun tiedot siirrettiin relaatiotietokantaan ja sieltä suoraan valmentajille valmiina pelaajien suorituskykyanalyysinä. Saksan menestys MM-kisoissa noteerattiin myös Wall Street Journal -lehdessä, joka nimitti big datan Saksan kahdenneksitoista kenttäpelaajaksi niissä kisoissa. (Mincev 2021, 29.)



Figure 4: SAP Germany's national team Snippet

Kuva 4. SAP:n HANA-järjestelmän sovellusnäkyminen ja analyysi joukkueen suorittamisesta MM-kisoissa 2014 (Mincev 2021, 29)

Englannin Valioliigassa pelaajien keskimääräinen rahallinen arvo on noussut 3,7 miljoonasta punnasta yli 12 miljoonaan puntaan kausien 2010–2019 aikana. Pelaajien suuri arvon nousu on ollut alkua ennustavan datalouhinnan (Predictive Data Mining) kehitykselle vammojen ennustamisessa. (Mincev 2021, 32.) Vammojen ennustamiseen käytetään puettavista teknologialaitteista saatua dataa. Saadakseen konkreettista tietoa vammojen ehkäisystä puettavien teknologialaitteiden kautta, on niistä kerätty data luokiteltu neljään pääryhmään: fysiologisiin muuttujiin, mekaanisiin kuormitusta mittaaviin muuttujiin, liikkumista mittaaviin muuttujiin sekä kiihtyvyy-, nopeus ja etäisyysmuuttujiin. (Mincev 2021, 32.) Teknologian ja IoT Sport-laitteiden kehittymisen myötä jalkapallossa on alettu investoida uusiin laitteisiin, jotta dataa saataisiin vieläkin suurempia määriä. Mincev kirjoittaa, että viime vuosien teknologian kehitys puettavassa tekniikassa on mahdollistanut urheilijoiden reaaliaikaisen liikkeen ja kuormituksen mittaamisen harjoituksen aikana. Tällaisen monipuolisen datan kerääminen on mahdollistanut ennustavan urheiluanalytiikan mallin. Nämä ennustavat mallit ovat dataan perustuvia päätöksentekomalleja, jotka pystyvät tunnistamaan ne henkilöt, joilla on suuri loukkaantumisen riski. (Mincev 2021, 33.)

Nopeudella ja pelinopeudella on suuri merkitys jalkapallossa. Yhden maailman arvostetuimman jalkapallovalmentajan Jose Mourinhon tunnetuimpia strategioita pelikentillä on ollut "bussin pysäköiminen" (Swartz 2017, 12). Termillä "bussin pysäköiminen" tarkoitetaan usein keskialueen tiivistä puolustuspelaamista, missä

puolustavan joukkueen pelaajat ovat lähes kaikki pallon ja oman maalin välissä. Tämän tarkoituksena Jose Mourinholla oli vaimentaa vastustajan joukkueen vauhteja heidän hyökätessään ja sitä kautta vähentää vastustajan saamia maalipaikkoja ja vastustajan tekemiä maaleja (Swartz 2017, 12). Pelin ja pelaajien vauhteja voidaan tutkia keräämällä dataa pelaajien seurantalaitteilla, kuten pelaajien käyttämällä siruilla tai puettavalla teknologialla. Tässä otetaan huomioon pelaajien nopeuden mittaaminen pituussuunnassa silloin, kun kyseisen pelaajan joukkueella on peliväline hallussaan. Tällä tavalla on esimerkiksi analysoitu, että nopeus vaikuttaa merkittävästi jalkapallossa ja tiedetään myös se, että maalintekoprosentti jalkapallossa kasvaa pelin edetessä, kun pelaajat alkavat väsymään. (Swartz 2017, 12.)



Image 3. The GPS kit is fairly standard: a single matchbook-sized sensor with vest and charging cable. Additional sensors such as foot pods or heart rate systems will be the new normal.

Kuva 5. Jalkapallossa käytettävä pelaajien puettava seurantalaitte, jolla voidaan esimerkiksi mitata pelaajien nopeuksia GPS:n avulla (Valle)

4.5 Data-analytiikan ja teknologian hyödyntäminen koripallossa

Koripallo on yksi eniten tilastoitu urheilulaji koko maailmassa. Vuorimies (2018, 7) kirjoittaa, että vuonna 2013 koripallossa, tarkemmin NBA:ssa, alettiin hyödyntämään big dataa. Ensimmäiset investoinnit NBA:ssa olivat vuonna 2014 asennetut SportVU-kamerat, joita asennettiin jokaisen NBA-joukkueen kotihalliin kuusi kappaletta. SportVU-kamerat

tallentavat lähes kaiken mitä kentällä tapahtuu, pallon liikkeestä pelaajien liikkeeseen ja jopa 25 datapistettä sekunnissa. (Vuorimies 2018, 7.)

SportVU-kameroiden tarkoitus urheiluanalytiikassa on täydentää tilastointia ja helpottaa datankeräämistä. Tilastointi on mekaanista työtä, ja koripallossa pelitilastoja ovat esimerkiksi pisteiden tekijät, levypallot, syöttömäärät ja heittomäärät. SportVU-kameroiden avulla tilastoista ja kerätystä datasta saadaan ymmärrys, joka kertoo, miksi asia kentällä tapahtui (Vuorimies 2018, 6–7). Esimerkiksi heiton mennessä koriin, on kameroiden avulla nähtävissä millainen heitto oli, mistä se tehtiin ja missä ajassa kaikki tämä tapahtui. Taustatutkimusten perusteella on havaittu, että pelaajien ja joukkueiden menestymiseen voidaan vaikuttaa tilastollisella mallinnuksella ja visuaalisilla järjestelmillä, kuten SportVU-kameroilla, lämpökartoilla, sosiaalisia verkostoja tutkimalla, biomekaniikan tilastoilla, harjoittelun mittaamisella sekä puettavilla tietotekniikkalaitteilla (Sarlis & Tjortjis 2020, 15). Sarlis ja Tjortjis (2020, 15) käyttävät artikkelissaan termiä Motion capture technologies. Motion capture -tekniikat keräävät valtavat määrät dataa pelaajista ja joukkueista yhden ottelun aikana. Näillä tekniikoilla pystytään seuraamaan pelaajien ja joukkueiden liikkeitä millisekuntin tarkkuudella. Motion capture -tekniikoilla tarkoitetaan teknologiaa, jonka avulla ottelun aikana pelaajista ja joukkueista kerätään dataa pelitilastoista, puettavista teknologialaitteista ja SportVU-kameroista. Kun nämä tekniikat ja niistä saatu data yhdistetään datalouhinnan ja koneoppimisen avulla, saadaan merkittäviä urheiluanalytiikan tuloksia, jotka vaikuttavat merkittävästi päätöksentekoon. (Sarlis & Tjortjis 2020, 15.)

Myös koripallossa niin kuin monessa muussakin urheilulajissa käytetään GPS-laitteita, joilla voidaan kerätä dataa pelaajien sijainneista ja liikkeistä pelin tai harjoitusten aikana. GPS-laitteet, joita pelaajat käyttävät ovat yleensä puettavia teknologialaitteita. (Vuorimies 2018, 7.) Näiden laitteiden avulla pelaajista saadaan seurantadataa, jolla voidaan analysoida pelaajien syöttöjä, liikkumista, arvioida heittojen riskiä tai syöttöjen laatua. Tällaisella datalla voidaan tehdä arvioita esimerkiksi yksittäisestä tilanteesta, kannattaako kolmenpisteenheittoa yrittää, jos omia pelaajia ei ole korin läheisyydessä vai kannattaisiko mieluummin syöttää tai ajaa itse korille. (Vuorimies 2018, 13–14.) Puettavat teknologialaitteet, kuten GPS-laite voivat kerätä dataa myös kiihtyvyyssantureiden, gyroskooppien, magnetometrien, sykemittareiden ja painesensoreiden avulla. Näiden ohjelmistot ovat yleensä ohjelmoituja C, C++, Aduino, Java, Swift tai Python ohjelmointikielillä. Kaikki tällaiset laitteet tuottavat esikäsiteltyä dataa älylaitteille, jotka ovat yleensä yhteydessä pilvipalveluihin. (Ryhänen 2020, 11-12.)

Sarlis ja Tjortjis (2020, 4) kuvaavat koripallonanalytiikkaa kuuden taulukon avulla. Taulukot kertovat miten pelaajista saatua dataa hyödynnetään koripallon eri osa-alueilla.

Ensimmäisessä taulukossa analysoidaan pelaajasta saatua dataa niin, että tarkastellaan pelaajan yleistä suorituskkyä $USG\%$:n (Usage Percentage) avulla, jolla voidaan mitata pelaajan sitoutumista peliin pelin aikana. Ensimmäisessä taulukossa on myös $NetRtg$ (Net Rating), joka laskee pelaajien pisteiden määrää verrattuna vastustajan pelaajiin, WS (Win Shares), jolla analysoidaan hyökkäyspeliä todella tarkasti sekä $eFG\%$ (Effective field goal percentage) ja $TS\%$ (True shooting percentage), jotka yhdessä selittävät pelaajien heittämisskyä. Viimeisenä ensimmäisessä taulukossa on $REB\%$ (Rebound percentage), joka laskee pelaajan voittamat levypallot ottelun aikana. (Sarlis & Tjortjis 2020, 4.)

Toisen taulukon avulla analysoidaan puolustuspeliä, joka sisältää pallonriistot ja pallon torjunnat, eli blokit. $DefRtg$:n (Defensive Rating) avulla analysoidaan pelaajan vaikutusta peliin puolustuspäässä siten, kuinka iso ero on vastustajan tekemien korien määrässä, kun pelaaja on kentällä ja kun hän on vaihdossa. Puolustuspeliin suuresti vaikuttavia tekijöitä, joita mitataan analyttisesti koripallossa ovat poikkeamat (Deflections) eli toisin sanoen tilanteet, joissa puolustava pelaaja osuu jollain ruumiinosallaan palloon tarkoituksellisesti tai vahingossa niin, että se muuttaa suuntaa ja pallo ei ole hetkeen kummallakaan joukkueella hallussa. Puolustuspelaamisessa käytetään termiä "really big plays", sellaisista suorituksista, joissa puolustava pelaaja estää vastustajan heiton ratkaisevalla hetkellä ja innostaa joukkuettaan tällä suorituksella. $BLK\%$ (Block percentage) taas laskee joukkueen blokkauksia, eli heittojen torjuntaja verrattuna vastustajan heittoyrityksiin. (Sarlis & Tjortjis 2020, 4.)

Table 2
Defensive criteria - Advanced basketball statistics.

Glossary	Description	Metric Type	Explanation
$BLK\%$	Block Percentage or Block Rate	Player	The opponent % two-point field goal attempts blocked while he was on the floor. $100 \cdot (BLK \cdot (TMP/5)) / (MP \cdot (OFGA - O3PA))$
$DREB\%$	Defensive Rebound Percentage	Player	The % of available defensive rebounds a player grabbed while he was on the floor.
$PF\%$	Percent of Team's Personal Fouls	Player	The % of a team's personal fouls that a player has while on the court
$STL\%$	Steal Percentage	Player, Team	The % of opponent possessions that end with a steal by the player on the floor. $100 \cdot (STL \cdot (TMP/5)) / (MP \cdot OP)$
BLK	Blocks	Player, Team	A block occurs when the defense player tips the ball, blocking their chance to score
Deflections	Deflections (Values of avg ranges from 0 to 5)	Player, Team	The number of the defense tackle the ball on a non-shot attempt
$DREB$	Defensive Rebounds	Player, Team	The number of rebounds a player or team has collected while they were on defense
STL	Steals	Player, Team	Number of times that takes the ball from a player on offense, causing a turnover
$DefRtg$	Defensive Rating	Player, Team	The number of points allowed per 100 possessions by a team. For a player, it is the number of points per 100 possessions that the team allows while that individual player is on the court. The formula is: $100 \cdot ((Opp\ Points) / (Opp\ POSS))$.
$DBPM$	Defensive Plus/Minus	Player, Team	A box score estimates of the defensive points per 100 possessions a player contributed above a league-average player, translated to an average team.
$DEF\ EFF$	Defensive Efficiency	Team	The number of points a team scores per 100 possessions. The formula is: $(100 \cdot Opp\ Points) / (Opponent\ FGA + Opponent\ TOV + (0.44 \cdot Opponent\ FTA) - Opponent\ OREB)$
DPR	Defensive Player Rating	Player	The formula is: $(Player\ spg + Player\ bpg / team\ minutes\ played) - (times\ blown\ by \cdot Pace\ of\ Players\ Era) \cdot Total\ Average\ of\ Possessions + (Players\ DRTG \cdot Team\ Pace) / Total\ number\ of\ years\ played$
$DRPM$	Defensive Real Plus Minus	Player, Team	Player's estimated on-court impact on team defensive performance, measured in points allowed per 100 defensive possessions.
DWS	Defensive Win Shares	Player	The number of wins contributed by a player due to his defense.

Kuva 6. Koripallon puolustuspelin analyysikäsitteet, mitkä ovat tekstissä kuvattu toisena taulukkona (Sarlis & Tjortis 2020, 9)

Kolmannen taulukon avulla analysoidaan kerättyä dataa hyökkäyspelin näkökulmasta. Hyökkäyspelissä tarkkaillaan pisteitä, levypalloja, syöttöjä ja heittoprosentteja. Myös hyökkäyspelissä analysoidaan pelaajien kykyä riistää pallo takaisin omalle joukkueelle, samalla tavalla kuin puolustuspelaamisessakin. Lisäksi hyökkäyspelaamisessa analysoidaan screenien tekemistä (Screen Assists metric), eli sitä kun oman joukkueen pelaaja estää sääntöjen puitteissa vastustajan pelaajaa liikkumasta palloa pitävän pelaajan eteen niin, että palloa pitävä pelaaja pääsee yrittämään koria. AST% (Assist percentage) mittaa sitä, kuinka monta pelaajan antamaa syöttöä on johtanut jonkun toisen pelaajan heittämään koriin pelinaikana, ottaen myös huomioon tämän pelaajan peliajan. Points per possession (PPP) mittaa pelaajan tehokkuutta, eli verrataan pelaajan tekemiä pisteitä hänen pallon hallussapitoaikaansa pelin aikana. (Sarlis & Tjortis 2020, 5.)

Table 3
Offensive criteria - Advanced basketball statistics.

Glossary	Description	Metric Type	Explanation
3PA%	3 Point Field Goals Percentage	Player, Team	The % of 3-point field goals attempted while on the court
3PM%	3 Point Field Goals Percentage Made	Player, Team	The % 3-point field goals made while on the court
BLKA%	Percent Blocked Field Goal Attempts	Player, Team	The % own blocked field goal attempts while on the court
FGA%	Field Goal Percentage Attempted	Player, Team	The % field goals attempted while on the court
FGM%	Field Goal Percentage Made	Player, Team	The % made field goals while on the court
FTA%	Free Throw Percentage Attempted	Player, Team	The % made free throws has made while on the court
FTM%	Free Throw Percentage Made	Player, Team	The % made free throws while on the court
FTr	Free Throw Factor	Player, Team	One of Four Factor. How often it goes to line and how often they made it. FT/FGA
FTM/FTA%	Percent of Team's Free Throws Made	Player, Team	Team free throw attempts made per field goal attempt
OREB%	Offensive Rebound Percentage	Player	The % of available offensive rebounds a player took while is on the floor
PF%	% of Team's Personal Fouls Drawn	Player, Team	The % of a team's personal fouls drawn by a player has while on the court
PTS%	% of Team's Points	Player, Team	The % of a team's points that a player has while on the court
PTS 2PT%	% of Points (2-Point Field Goals)	Player, Team	The % of points scored by a player or team from 2-point field goals
PTS 3PT%	% of Points (3-Point Field Goals)	Player, Team	The % of points scored by a player or team from 3-point field goals
PTS FBPS%	% of Points (Fast Break Points)	Player, Team	The % of scored points by a player or team from fast break opportunities
PTS FT%	Percent of Points (Free Throws)	Player, Team	The % of scored points by a player or team from free throws
2nd PTS	Second Chance Points	Player, Team	The % of isolation plays that shoots free throws of a shooting foul
3PA	3 Point Field Goals Attempted	Player, Team	The number of 3-point field goals that a player or team has attempted
3PM	3 Point Field Goals Made	Player, Team	The % of a team's 3-point field goals made while on the court
FBPS	Fast Break Points	Player, Team	The number of points scored by a player or team while on a fast break
FGA	Field Goals Attempted	Player, Team	The number of 2-point field goals attempted
FGA/ Poss	FGA/Possession	Player, Team	Calculated the shot attempts in each possession.
FGM	Field Goals Made	Player, Team	The number of 2-point field goals made
FTA	Free Throws Attempted	Player, Team	The number of free throws attempted
FTM	Free Throws Made	Player, Team	The number of free throws made
OR/P	Offensive Rebounds/Possession	Player, Team	Offensive Rebounds per completed possession
OREB	Offensive Rebounds	Player, Team	The number of rebounds gathered while they were on offense
PF	Personal Fouls Drawn	Player, Team	The number of personal fouls that are drawn by a player or team
PTIP	Points in the Paint	Player, Team	The number of points scored by a player or team in the paint
PTS/Poss	Points/Possession	Player, Team	The made points in each time touches the ball.
PTS	Points	Player, Team	The number of scored points.
PTS Off	Points Off Turnovers	Player, Team	The number of points scored following an opponent's turnover.
Tov			
OBPM	Offensive Plus/Minus	Player	A box score estimates of the offensive points per 100 possessions a player
OFF EFF	Offensive Efficiency	Team	The number of points a team scores per 100 possessions. The formula is: $(100 * \text{Points} / (\text{Team FGA} + \text{Team TOV} + (0.44 * \text{Team FTA}) - \text{Team OREB}))$
OffRtg	Offensive Rating	Player, Team	Measures a team's or player (on court) points scored per 100 possessions. $(100 * \text{Points} / (\text{Team FGA} + \text{Team TOV} + (0.44 * \text{Team FTA}) - \text{Team OREB}))$
ORPM	Offensive RPM	Player, Team	Player's on-court impact on team offensive performance in points scored per 100 offensive possessions
OWS	Offensive Win Shares	Player	The number of wins contributed by a player due to offense.
PPP	Points Per Possession (Values of avg ranges up to 1.8)	Player, Team	The number of points a player or team scores per possession $\text{PTS} / (\text{FGA} + 0.44 * \text{FTA} + \text{TOV})$

Kuva 7. Koripallon hyökkäyspelin analyysikäsitteet, mitkä ovat tekstissä kuvattu kolmantena taulukkona (Sarlis & Tjortjis 2020, 10)

Neljättä taulukkoa Sarlis ja Tjortjis (2020, 5) kuvaavat yleiseksi koripalloanalyysikäsitteiden taulukoksi, jossa käsitteet ja urheiluanalytiikan kuvaukset ovat hyvin yleisiä joukkueurheilulajeille. TOV% (Turnover percentage) keskittyy prosentiosuuteen, joka lasketaan pelaajan pallon menetyksien määrä verrattuna pelaajan peliaikaan (Sarlis & Tjortjis 2020, 5). Aikaisemmat taulukot kuvasivat yksittäisten pelaajien tekemiä suorituksia kentällä, mutta joukkuelajeissa ja urheiluanalytiikan kannalta on otettava huomioon myös joukkuepelaaminen. Neljännessä taulukossa yhdistetään edellisten taulukoiden analyysimenetelmiä, jonka avulla pyritään löytämään enemmän

joukkuepeleä edistävää analytiikkaa. AST/FGM on yhdistetty Assist percentage -tilasto ja Field goal made -tilasto. Näiden kahden tilaston yhdistämisen avulla pystytään tarkastelemaan ja analysoimaan, kuinka monta kertaa kaikista oman joukkueen pelaajien antamista syötöistä on tehty koreja pelin aikana. (Sarlis & Tjortjis 2020, 5.) Tällä analyysimenetelmällä voidaan tarkastella esimerkiksi sitä, millainen syöttöpeli joukkueella on; johtaako syötöt yleensä koreihin vai onko syöttämisessä parantamisen varaa.

Table 4
Overall Performance criteria - Advanced basketball statistics.

Glossary	Description	Metric Type	Explanation
AST%	Assist Percentage	Player, Team	The % of teammate's FG that a player assists when is on the floor $100 * AST / (((MP / (TMP / 5)) * TFG) - FG)$ AST = Assists, MP = Minutes Played, TMP = Team Minutes Played, TFG = Team Field Goals, FG = Field Goals
TOV%	Turnover Percentage	Player, Team	The number of turnovers committed per 100 possessions. One of the recognized Four Factors. $100 * TOV / (FGA + 0.44 * FTA + TOV)$
TRB%	Total Rebound Percentage	Player	$100 * (TRB * (TMP / 5)) / (MP * (TRB + OTRB))$. It is a weighted average of total rebounds a player took while he was on the floor.
AST/Poss	Assists per Possession	Player, Team	It shows how well passed the ball through the game.
AST	Assists	Player, Team	The number of assists that goes to a made basket
BLKA	Blocks against	Player, Team	The number of shots attempted and blocked by a defender
DD2	Double doubles	Player	The number of double-digit number total in two of the five statistical categories in a game that a player achieves
TD3	Triple doubles	Player	The number of double-digit number total in three of the five statistical categories in a game that a player achieves
FT/ Poss	FT/Possession	Player, Team	The free throw shot attempt per every possession.
FTA RATE	Free Throw Attempt	Player, Team	The number of free throws attempts in comparison to the number of field goal attempts
GP	Games Played	Player, Team	The number of games a team or player played
L	Losses	Player, Team	The number of games lost by a team or player
MPG/MIN	Minutes Played	Player	The number of minutes played by a team or player
PF	Personal Fouls	Player, Team	The number of personal fouls a player or team committed
PRL	Position Replacement Level	Player, Team	PRL = 11.5 for power forwards, 11.0 for point guards, 10.6 for centers, 10.5 for shooting guards and small forwards
REB	Rebounds	Player, Team	The number of total rebounds a team or player has collected on either offense or defense
STL/DP	Steals/Defensive Possession	Player, Team	How many steals your defense gets for every one of your opponent's offensive possessions.
TOV/Poss	Turnovers/Possession	Player, Team	How often a team or player made a turnover every time you touch the ball.
TOV	Turnovers	Player, Team	A turnover occurs when the player or team on offense loses the ball to the defense
W	Wins	Player, Team	The number of games won by a team or player
%WIN	Win Percentage	Player, Team	W/GP. The percentage of games played that a player or team has won
Loose Ball Rec	Loose Ball Recovered (Values of avg ranges from 0 to 12)	Player, Team	The defensive or offensive actions while trying to secure a loose ball
AST/TOV	Assist to Turnover Ratio (Values of avg ranges from 0 to 3)	Player, Team	The number of assists for a player or team compared to the number of turnovers they have committed

Kuva 8. Koripallon analyysikäsitteet, mitkä ovat tekstissä kuvattu neljäntenä taulukkona (Sarlis & Tjortjis 2020, 11)

Viides taulukko osoittaa edistyneen koripalloanalytiikan vertailumatriiseja. Tämä tarkoittaa sitä, että jos urheiluanalytiikassa käytettäisiin aina vain laskentamallia, jolla suuremman arvon saanut joukkue tai pelaajat olisivat voittajia, ei urheiluanalytiikasta olisi niin paljon hyötyä. (Sarlis & Tjortjis 2020, 5.)

Kuudes taulukko osoittaa urheiluanalytiikan tutkimusten avulla bibliografisia- ja luokiteltuja matriisimalleja, joissa käytetään data mining (DM) algoritmeja ja tekniikoita (Sarlis & Tjortjis 2020, 5). Sarlis ja Tjortis (2020, 5) kirjoittavat artikkelissaan, että taulukko kuusi ja aiheesta tehdyt kirjallisuustutkimukset todistavat sen, että urheiluanalytiikka on kehittyvä

tieteenala, joka syvenee DS- (Data structure), ML-, ja DM-tekniikoita käyttämällä suorituskyvyn optimoimiseksi ja ennustamiseksi. Tulevaisuudessa edellä mainitut tekniikat tulevat kehittymään ja kehittämään urheiluanalytiikkaa entisestään.

Koripallo on joukkuelaji, eikä analytiikan merkitys ole vain tehokkaiden pelaajien tai joukkueiden erottamisessa (Sarlis & Tjortjis 2020, 14). Koripallossa jokaisessa ottelussa yhdellä joukkueella on kokoonpanossa kaksitoista pelaajaa. Analytiikan kehittymisen myötä on pystytty kehittämään koripalloa ja pelaajia niin, että on pystytty takaamaan kaksitoista kelpoista pelaajaa. Datan avulla on myös huomattu, että vaihtopelaajilla on suuri merkitys pelien lopputuloksiin. Uuteen vaikuttavaan rooliin ovat nousseet joukkueen roolijakautuminen ja oikeatasapaino joukkueen sisällä. Ero tässä lähestymistavassa on muuttunut huomattavasti edellisenä vuosikymmenenä. (Sarlis & Tjortjis 2020, 14.)

Kun puhutaan koripallosta ja urheiluanalytiikasta, on nostettava esiin tapaus Stephen Curry. Stephen Curry on NBA-joukkueen Golden State Warriorsin pelintekijä ja takamies, (point guard) jolla on pituutta vain 188 senttimetriä, eli hän on lyhyt NBA-pelaajaksi (NBA.com 2021). Vuonna 2017 Stephen Curry rikkoi sarjan ennätyksiä kolmenpisteen heitoilla. Ennen tätä NBA oli pitkien ja isojen miesten peli, mutta Curryn voitettua NBA:n mestaruus ja MVP-palkinto (Most Valuable Player), alkoivat muut joukkueet tutkia kolmenpisteen heiton strategiaa. NBA:ssa alettiin puhumaan, että Stephen Curry muutti kaiken sen, miten koripalloa on ennen pelattu. Vuonna 1998 joukkueet yrittivät kolmen pisteen heittoja keskimäärin 12,7 kertaa pelin aikana, kun vuonna 2019 kolmen pisteen heittoja yritetään keskimäärin 31,3 kertaa pelin aikana. (Goldsberry 2019.) Kun tiedetään edellä mainittu esimerkki Currysta ja se kuinka paljon dataa on saatavilla huippu-urheilussa, voidaan päätellä, että datan arvo on noussut viime vuosina selkeästi. Tästä syystä on ymmärrettävää, että joukkueet eivät välttämättä halua paljastaa kaikkia mahdollisia dataan liittyviä tekniikoita muille joukkueille.

4.6 Data-analytiikan ja teknologian hyödyntäminen NFL:ssä

Viime vuosina NFL, eli National Football League on ryhtynyt kehittämään liigassa pelaavia joukkueita data-analytiikkaa hyödyntäen Next Gen Stats -projektissaan. Vuonna 2014 NFL upotti seurantasirut kaikkien pelaajien hartiasuojiiin, jolloin liiga pystyi seurata kunkin pelaajan sijaintia, nopeutta ja suuntaa pelin aikana. (Avery 2021, 2.) Kolme vuotta myöhemmin vuonna 2017 NFL lisäsi samanlaiset seurantasirut pelivälineeseen, eli palloon, ja vuoteen 2018 mennessä NFL julkaisi kaikki nämä seurantatiedot jokaiselle joukkueelle niin, että joukkueet pystyivät itsekkin hyödyntämään peleistä ja pelaajista kerättyä dataa. Kun NFL oli saanut joukkueet mukaan projektiinsa datan keräämisessä, antoi se osan näistä tiedoista yleisön saataville vuonna 2019. Nämä tietotekniikan ja datankeräämisen edistysaskeleet ovat vieneet NFL:n data-analytiikan vallankumoukseen,

ja monet NFL-joukkueet käyttävät nyt tilastoja ja koneoppimisen asiantuntijoita saadakseen kaiken mahdollisen edun kentällä. (Avery 2021, 2.)

Loukkaantumisriski NFL:ssä on erittäin suuri, koska amerikkalainen jalkapallo on fyysinen laji ja kontakteja kentällä tulee paljon. NFL on tehnyt yli neljäkymmentä sääntömuutosta välttääkseen päähän kohdistuneita loukkaantumisia onnistumatta. Sääntömuutosten täytäntöönpanon aikana aivotärähdysten määrä saavutti kaikkien aikojen suurimman määrän kaudella 2017, jolloin tapahtui 281 aivotärähdystä. Kaudella 2018 aivotärähdysten määrä tippui 214 kappaleeseen, mutta nousi taas kaudella 2019 224 kappaleeseen. (Ghasem, Saxon & Valenzuela 2021, 436.)

Pelkät aivotärähdykset eivät ole NFL:n ainoa loukkaantumisongelma. Tuki- ja liikuntaelimistön sekä pehmytkudosten vammat ovat niin yleisiä, että niistä kärsivät uransa aikana lähes jokainen pelaaja. Pelkästään alaraajavammat aiheuttavat 60 % väliin jääneistä peleistä koko sarjassa, ja niitä esiintyy yhteensä noin kuusisataa kertaa kauden aikana. Näistä vammoista yleisin on ACL-vamma (anterior cruciate ligament), joka on siis polven eturistisiteen repeämä. ACL-vammoja raportoitiin 109 pelaajalla kaudella 2019. Kun tulokset pelaajien vammoista otetaan huomioon, on selvää, että tuki- ja liikuntaelimistön vammojen ehkäisystä on tullut ykkösprioriteetti liigassa. (Ghasem, Saxon & Valenzuela 2021, 436.)

Vuonna 2017 NFL:n komissaari Roger Goodell kertoi julkisesti liigan tehneen yhteistyösopimuksen yhtiön Amazon Web Servicesin kanssa. Yhteistyösopimus takaa uuden teknologian ja datan käytön pelaajien turvaamiseen. On tutkittu, että amerikkalaisessa jalkapallossa tapahtuu merkittävästi vammoja -jopa aivovammoja- jonka seurauksena liiga on valmis hyödyntämään uutta tekniikkaa. (Ghasem, Saxon & Valenzuela 2021, 436.)

NFL:n yhteistyö Amazon Web Servicesin kanssa laajensi huomattavasti saadun datan määrää ja analyysimahdollisuuksia. Next Gen Stats -projektissa on käytetty yli seitsemän vuotta järjestelmää, jonka on kehittänyt Zebra Technologies -yritys. Zebra Technologies tarjoaa NFL:lle radiotaajuustekniikkaa, (Radiofrequency Identification) joka varastoi pelaajien seurantadataa ja liikkumista kentällä. (Ghasem, Saxon & Valenzuela 2021, 436.) Pelaajien seurantadataa kerätään pelaajien hartiasuojissa olevilla siruilla, jotka ovat yhteydessä 20–30 laajakaistaiseen vastaanottimeen, jotka ovat sijoitettuna stadioneille. Anturit keräävät pelaajista pelaajien nopeutta, kiihtyvyyttä ja kokonaismatkaa. Nämä kerätyt tiedot ovat verrattavissa epäsuoriin mittareihin, jotka taas mittaavat pelaajien intensiteettiä ja tehokkuutta samalla, kun seurataan pelaajien sijaintia tuumissa jopa kymmenen kertaa sekunnissa. Pelaajista kerättyä dataa annetaan valmentajille ja faneille pelin taktisten oivallusten saamiseksi. (Ghasem, Saxon & Valenzuela 2021, 436.)

Amazon Web Servicesin uuden teknologian ja analyysien lupaus vammojen ehkäisyyn liittyy kykyyn löytää yhteys yksittäisten pelaajien kuormitusten, pelin intensiteetin sekä pelien määrän välillä; tämä mahdollistaa riskien tunnistamisen varhaisessa vaiheessa ja sitä kautta vammojen ehkäisemisen. Eniten vammoja esiintyy niillä pelaajilla, joilla on ollut suuri harjoittelukuormitus tai nopea työmäärän kasvaminen. Monet yksilön ominaisuudet, kuten ikä, harjoitusintensiteetti, loukkaantumishistoria ja fyysinen kyky yhdistettynä harjoittelukuormitukseen, ovat kaikki tekijöitä, jotka altistavat loukkaantumisriskille. (Ghasem, Saxon & Valenzuela 2021, 437.) Ghasem, Saxon ja Valenzuela (2021, 437) nostavat esiin artikkelissaan selvän ongelmakohdan pelaajien seurantalaitteiden käytöstä ja harjoittelun kuormituksesta. Pelaajien seurantalaitteita ei käytetä harjoittelukausilla, joka tarkoittaa sitä, että pelaajien harjoittelukuormituksesta ei saada kokonaisvaltaista kuvaa. Loukkaantumisriski nousee, jos pelaaja ei kerkeä päästä ennen kauden alkua täyteen kuntoon tai vastaavasti ylikuormittuu jo harjoituskauden aikana.

Table.
List of wearable devices approved for use in NFL games (Zebra) or practice.

Company	Product	Product Form	Primary Product Functionality
Zebra Technologies	Zebra MotionWorks Sport	Device chips in uniforms, pads, and footballs connected to sensors and receivers throughout stadiums.	RFID ^a used to quantify player and football movement and distance profiles.
Catapult Sports	OptimEye S5	Wearable device in upper body garment	GPS ^b and GLONASS ^c antenna tracking of sport-specific movement, player load, orientation, and heart rate.
KINEXON	KINEXON; SafeTag	Wearable device in upper body garment; Wearable wrist sensor for contact tracing	UWB ^d designed to record player position, motion, orientation, and heart rate; UWB measuring physical distance and length of time between users.
TITAN	TITAN 2 +	Wearable device in upper body garment	Quadruple GLONASS aimed for player tracking and inertial sensors quantifying performance block.
Polar	Polar Team Pro	Wearable device in upper body garment or around trunk	GPS and MEMS ^e designed to measure performance metrics and heart rate.
STATSports Group Limited	Apex	Wearable device in upper body garment	GPS designed to monitor player load and performance, collision analysis, and heart rate.

^aA wireless system comprised of two components: tags and readers. The reader is a device that has one or more antennas that emit radio waves and receive signals back from the RFID tag. Tags, which use radio waves to communicate their identity and other information to nearby readers, can be passive or active.

^bA navigational system using satellite signals to fix the location of a radio receiver on or above the earth's surface.

^cA satellite-based radio navigation system run by the Russian Ministry of Defense. It uses 21 medium earth orbit (MEO) satellites and three spares. Similar to the GPS in the United States, GLONASS enables 3D positioning anywhere on earth within 100 m to 150 m for the public and 10 m to 20 m for the military.

^dA radio-based communication technology for short-range use for the fast and stable transmission of data indoors and outdoors.

^eTechnology defined as miniaturized mechanical and electro-mechanical elements (*i.e.*, devices and structures) that are made using the techniques of microfabrication. Microsensors are as transducers, which are devices that convert energy from one form to another. In the case of microsensors, the device typically converts a measured mechanical signal into an electrical signal for almost any sensing modality.

Kuva 9. NFL-liigan hyväksymät teknologialaitteet pelin aikana (Ghasem, Saxon & Valenzuela 2021, 437)

NFL ja Amazon Web Services ovat työskennelleet yhdessä luodakseen Digital Athlete -alustan, jota kuvataan artikkelissa tietokonesimulaatiomalliksi, jossa voidaan käydä läpi loputtomasti erilaisia pelissä tapahtuvia skenaarioita peliympäristössä ilman pelaajan loukkaantumisriskiä. Amazon tekee yhteistyötä myös joukkueiden kanssa. Amazon teki esimerkiksi sopimuksen NFL:ssä pelaavan joukkueen Seattle Seahawksin kanssa pilvi-, koneoppimis- ja tekoälypalveluista. Yhteistyö Amazonin kanssa auttaa Seahawksia

seuraamaan pelaajien suorituskykyä sekä peleissä, että harjoituksissa keräämällä dataa pelaajien ja pallon liikkeistä. Kun data on kerätty ja analysoitu, Amazon Web Services voi rakentaa ennustavia malleja pelitapahtumista. Ennustavista malleja voidaan rakentaa tuhansia erilaisia ja niistä tehdään skenaarioita pelissä tapahtuvista asioista; niillä voidaan tutkia esimerkiksi pelintilanteita, kuten miten rynnätä kohti vastustajan pelinrakentajaa parhaiten, löytää toimivimmat reitit hyökkäyspeliin, puolustuspelin linjan tiiviyyttä ja toimintaa erikoisjoukkueiden ollessa pelikentällä, kokoonpanoja ja ottelumenestystä. (Ghasem, Saxon & Valenzuela 2021, 438.)

Edellä mainittuja tekniikoita on eniten käytetty fanien viihdyttämiseen ja sitouttamiseen ottelutapahtumiin sekä urheilumerialle. Esimerkiksi median mainoskatkojen aikana faneille näytetään tarkasti videolta dataa hyödyksi käyttäen, kuinka kovaa pelinrakentaja on juossut neljänneksen aikana tai kuinka puolustava pelaaja on tehnyt hienon pallonriiston. (Ghasem, Saxon & Valenzuela 2021, 438.) Data-analytiikan hyödyntäminen NFL:sta alkoi sillä, että vuonna 2014 NFL salli datan keräämisen pelaajista ja pelitapahtumista, ja vuonna 2017 liiga ryhtyi hyödyntämään dataa myös loukkaantumisten ehkäisyyn. NFL on tehnyt vuonna 2020 vuoteen 2030 kestävästä työehtosopimuksen NFLPA:n (National Football League Players Association) kanssa siitä, että pelaajien tulee käyttää seurantasensoreita aina peleissä. (Ghasem, Saxon & Valenzuela 2021, 438.) Kuvaa hyvin urheiluanalytiikan kehitystä, että seitsemän vuotta sitten NFL:ssä joukkueet eivät saaneet käyttää seurantalaitteita ja tänä päivänä niitä on pakko käyttää.

5 Pohdinta

5.1 Yhteenveto

Tässä tutkielmassa lähdettiin hakemaan vastausta seuraaviin kysymyksiin: minkälaisia datankeräysmenetelmiä, data-analytiikkaa ja teknologiaa käytetään erilaisissa urheilulajeissa? Miten urheilussa kerätty data vaikuttaa urheilijoihin, joukkueisiin ja itse urheiluun? Mihin urheilussa kerättyä dataa käytetään? Tutkielman tavoitteena oli muodostaa eheä näkemys datankeräämisestä, data-analytiikasta ja teknologian hyödyntämisestä huippu-urheilussa.

Kuvailevan kirjallisuuskatsauksen menetelmällä tutkielma onnistuin luomaan ymmärrettävän sekä ajankohtaisen tilannekatsauksen urheilujoukkueiden, urheilijoiden sekä urheilusarjojen datan hyödyntämisestä heidän toiminnassaan. Aihetta on aiemminkin tutkittu, mutta aiheen ollessa ajankohtainen ja uusi, on lisätutkimukselle tulevaisuudessa väistämättä tarvetta.

Yhteen kysymykseen tutkielma ei löytänyt suoraa vastausta: ”Miten urheilussa kerätty data vaikuttaa urheilijoihin?” On vaikeaa tutkia, miten urheilijasta kerätty data suoranaisesti vaikuttaa urheilijaan, sillä urheilijoiden kokemukset ja oppiminen on subjektiivista. Tähän kysymykseen eivät aikaisemmatkaan tutkimukset vastanneet suoranaisesti. Ainoana esimerkkinä aikaisemmissa tutkimuksissa nousi esiin urheilijan loukkaantumisen ja rasisuuden tarkastelu urheilijaan vaikuttavana tekijänä. Urheilijan loukkaantumisen tarkasteltiin erityisesti NFL:ssä, mutta se otettiin myös huomioon jalkapallosarjassa Valioliigassa. NFL-organisaatio ja sarjan joukkueet ovat sitoutuneet keräämään dataa sekä tutkimaan pelaajien loukkaantumisia ja loukkaantumisriskejä yhteistoimin. Niin jalkapallossa kuin amerikkalaisessa jalkapallossa dataa hyödynnetään loukkaantumisen kartoittamisessa pelaajien seurantalaitteilla, joilla mitataan pelaajien fyysistä suorituskykyä. Aikaisemmissa tutkimuksissa nousi esiin, että datan avulla on pystytty todentamaan yksittäisen urheilijan loukkaantumisen kasvaminen suhteessa urheilijan kovaan pitkäaikaiseen rasitukseen.

Aikaisemmat tutkimukset data-analytiikan ja teknologian hyödyntämisestä huippu-urheilussa käsitelivät aihetta pinnallisesti ja yleisesti. Aikaisempia tutkimuksia tarkastellessa on huomioitavaa, että urheilumaailma on hyvin tarkka siitä, mitä taktiikoista, pelaajista tai joukkueen sisäisistä tavoista kerrotaan pukukoppien ja kenttien ulkopuolelle. Aikaisempien tutkimusten perusteella on hyvin todennäköistä, että datankeräämisen,

data-analytiikan ja teknologian hyödyntämisen keinoja huippu-urheilussa pidetään tänä päivänä paljon joukkueiden sisäisenä tietona, eikä niitä haluta levittää. On mahdollista, että datasta on tulossa huippu-urheilussa niin arvokasta, että sitä ei avokätisesti haluta levittää kaikille. Tämä voi olla yksi syy hieman pinnallisiksi jääneisiin aikaisempiin tutkimuksiin.

Tutkielmassa nousi esille, että sarjat NFL, NBA ja NHL hyödyntävät kaikki hyvin samanlaista teknologiaa datan keräämisessä ja sen analysoimiseen. Tänä päivänä jokaisessa edellä mainitussa sarjassa hyödynnetään videokamerateknologiaa datan keräämiseen pelaajista ja joukkueista. Yksi videokamerateknologia on SporVU-kamerat ja niitä on käytössä NFL:ssä, NBA:ssa ja osassa jalkapalloseuroista. Liigassa käytetään kameroiden sijaan sensoriteknologiaa, joka on suomalaisen yrityksen Wisehockey Oy:n järjestelmä ja hyvin samantapainen SportVU-kameroiden kanssa. NHL:ssä käytetään SAP:n IoTSports -tekniikkaa, joka on myös hyvin samantapainen edellä mainittujen videokamerajärjestelmien kanssa. Kamerat tai sensorit asennetaan kiertämään hallia tai kenttään siten, että kentältä nähdään kaikki mahdolliset kulmat. Koripallokentillä NBA:ssa SportVU-kameroita on kuusi ja jalkapallo kentällä niitä voi olla jopa kahdestakymmenestä neljästä kamerasta seitsemäänkymmeneen kameraan. On huomioitavaa, että kaikki suurimmat urheilusarjat hyödyntävät hyvin samanlaisia datankeräysmenetelmiä. Huippu-urheilussa ei siis ainoastaan tarkkailla oman urheilulajin kilpailijoita ja heidän toimintatapojaan, vaan voidaan mallia ottaa myös erilaisista urheilulajeista. Tämä tukee aikaisemmassa kappaleessa mainittua päätelmää datan arvon nousemisesta, mutta myös tämä tutkielman on tarpeellisuudesta sekä hyödynnettävyydestä, koska data ja teknologiat, joita eri urheilulajeissa hyödynnetään ovat hyvin samanlaisia urheilulajista riippumatta.

Kameroiden lisäksi kaikissa urheilulajeissa pelaajat käyttävät siruteknologiaa. Siruteknologia on kaikissa tutkielman lajeissa hyvin samanlaista ja ne sisältävät GPS-paikantimet, joiden avulla on mahdollista seurata pelaajan liikkeitä pelikentällä pelinaikana tai harjoituksissa. Osa urheilulajeista hyödyntää siruteknologiaa hieman monipuolisemmin kuin toiset, mutta pääsääntöisesti niistä kerätään samanlaista dataa urheilulajista riippumatta. Sirujen avulla pelaajista mitataan esimerkiksi pelaajan nopeutta, matkaa ja liikkumista kentän eri osa-alueilla. Kaikissa tutkielman lajeissa maalit tai korit ovat pelin pääroolissa, joten siruteknologialla voidaan esimerkiksi tarkastella sitä, kuka teki maalin ja mistä paikasta se tehtiin. Hieman kehittyneemmässä urheiluanalytiikassa voidaan tarkastella esimerkiksi sitä, kun tehdään maali, missä paikassa kukin pelaaja on ollut maalin aikana ja kenen liikkeet ovat edesauttaneet siihen, että peliväline on mennyt maaliin. Ja tietysti myös päinvastoin, kun tulee takaisku eli peliväline menee omaan

maaliin, voidaan tarkastella puolustuksen ryhmittymistä ja löytää syy seuraussuhteita siitä, miksi vastustaja onnistui maalinteossa.

Seurantasirujen lisäksi pelaajilla on päällään usein puettavia sensorilaitteita, joilla pystytään seuraamaan heidän fyysisiä ominaisuuksiansa, kuten sydämensykkettä. NFL erottui kaikista muista tutkielman urheilulajeista siinä, että NFL-organisaatio ja joukkueet ovat ottaneet sensorilaitteet erityisesti käyttöön loukkaantumisriskien havaitsemisessa. On hyödyllistä ymmärtää pelaajien rasitusta myös valmentajan näkökulmasta. Esimerkiksi NHL:ssä ja NBA:ssa pelataan yli kahdeksankymmenen pelin runkosarja ja se on erittäin raskas pelaajille. Pelaajien rasituksen seuraaminen vähentää loukkaantumisriskiä, mutta mahdollisesti myös tehostaa joukkueen pelaamista, koska valmentajalla on mahdollisuus valita energisimmät pelaajat kentälle joka peliin.

Teknologian kehittyminen on mahdollistanut monen pelitapahtuman seuraamista videolta koneoppimisen kautta. Koneoppimisella voidaan luoda pelitapahtumista skenaarioita ja tarkastella esimerkiksi toimivia taktiikoita hyökkäyspeliin. Koneet tekee tuhansia eri skenaarioita ja skenaarioiden lopputuloksia tarkastelemalla on mahdollista löytää toimiviakin ratkaisuja pelitilanteisiin. Menemättä syvemmin ennustaviin koneoppimisen malleihin, tutkielmassa esiintyi tekoälyn hyödyntämistä datankeräämisestä videokuvasta niin, että pelivälineissä ja pelaajissa olevat sirut rekisteröivät erilaisia pelitapahtumia suoraan järjestelmiin, jotka luovat datasta valmista urheiluanalytiikkaa. Esimerkiksi urheilussa ei enää tarvitse tilastoida pallon- tai kiekonmenetyksiä mekaanisesti, vaan SportVU-kamerat lukevat sirun pelivälineestä ja pelaajasta, joka pelivälinettä kuljettaa ja ymmärtää pelitapahtumista sen, koska peliväline vaihtaa omistajaa niin, että järjestelmä kirjaa sen pallon- tai kiekonriistoksi. Tekoälyn hyödyntämisellä datan käytöstä tulee entistä taktisempaa ja niin valmentajat kuin pelaajatkin pystyvät hyödyntämään sitä erilaisissa tilanteissa kehittääkseen peliä. Tutkielmassa nousi esiin, että koneoppimista ja tekoälyä hyödynnetään tällä hetkellä urheilussa enemmän vedonlyöntiin kuin itse urheiluun. Poikkeuksena tutkielmassa nousi jälleen esiin NFL ja joukkue nimeltä Seattle Seahawks, joka oli hyödyntänyt tekoälyä ja luonut hyökkäyspään taktiikkaa mallintavia skenaarioita koneoppimista hyödyntäen. Näin pelaajien ei tarvinnut rasittua ja harjoitella hyökkäyspään kuvioita vaan niitä pystyttiin syöttämään väsymättömälle tietokoneelle. Tämän tarkoituksena oli löytää toimivia kuvioita erilaisiin hyökkäyspelin pelitilanteisiin.

5.2 Tutkimuksen eettisyys sekä luotettavuus

Tämä tutkielma ei ole kohdannut toteutuksessaan eettisiä ongelmia, sillä tutkielma toteutettiin kuvailevana kirjallisuuskatsauksena, eikä esimerkiksi empiirisesti haastatteluna. Lähdeaineiston ollessa validia sekä reliaabelia on mahdollista luottaa aikaisempien tieteentekijöiden harkintaan eettisyydestä, joten tässä tutkimuksessa ei ole

tarvinnut keskittyä esimerkiksi tutkimuslupien saantiin taikka tutkimushenkilöiden henkilöllisyyden salaamiseen.

Tutkimuksen luotettavuutta on mahdollista tarkastella kahdesta eri näkökulmasta, jotka ovat tämän tutkimuksen luotettavuutta lisäävät sekä heikentävät tekijät. Ensin käsitellään luotettavuutta lisäävät faktorit. Läpi tutkielman raportointi sekä viittaukset ovat olleet avoimia, ja lähteinä on käytetty vain luotettavia tieteellisiä- taikka asiantuntijalähteitä. Lähdeaineiston rajaaminen on olennaista, etenkin kun löydettävissä on monen tasoista lähdettä. Yksi luotettavuutta lisäävistä tekijöistä onkin laaja lähdeaineisto, joka mahdollisti aineiston keskinäisen ja sisäisen vertailun. Laajan aineiston pohjalta ei tarvitse luottaa vain yhden aikaisemman tieteen tekijän sanaan, vaan pystyi lähteen luotettavuutta lisäämään, kun sama ”lopputulos” oli löydettävissä monesta muustakin lähteestä. Tämä kuvaileva kirjallisuuskatsaus on toteutettu siis luotettavalla, mutta myös ajankohtaisella materiaalilla. Kuten aikaisemmin nostettu esille; teknologian ollessa tutkimuskohteena, on tärkeää keskittyä ajankohtaiseen lähdeaineistoon ja rajata ulkopuolelle vanhentunut tieto. Tähän seikkaan käytettiin erityisesti huomiota, sillä tutkielman fokus on keskittyä huippu-urheilun nykyaikaan. Koko tutkielman läpi on löydettävissä punainen lanka, jonka avulla tutkimuskysymyksiin vastataan teoreettisesti. Tutkielman toteutus on ollut tarkkaa, lähteet ovat luotettavia, tulosten tarkastelu avointa sekä tehdyt analyysitoimenpiteet perusteltuja.

Tutkielman luotettavuutta mahdollisesti laskevassa mielessä on tarkasteltava kielirajauksen kannalta. Moni tässä opinnäytetyössä käytetyistä lähteistä oli alkuperäiskieleltään englanniksi, mutta tämän työn toteutuskielen ollessa suomi, olivat käännökset välttämättömiä paikoitellen. Kun kyseessä on käänнос kieleltä toiselle, on mukana aina riski käännosvirheille, ja näitä pyrittiinkin torjumaan huolellisella valmistelulla sekä kielellisen tarkkuuden vaalimisella. Kokonaisuutena tutkielma on onnistunut vaalimaan tieteen eettistä ohjenuoraa kuin myös luomaan eheän ja validin kokonaisuuden hyödyntäen luotettavia ja ajankohtaisia lähteitä. Tutkimuskysymyksiin sekä -tavoitteisiin päästiin onnistuneesti, ja tämä kirjallisuuskatsaus on kohdennettavissa edelleen samalle lukijakunnalle, kun alkuperäinen tarkoitus olikin.

5.3 Mahdolliset jatkotutkimukset

Kuten yhteenveto kappaleessa mainittiin, aikaisemmat tutkimukset eivät vastanneet tutkimuskysymykseen: ”Miten urheilussa kerätty data vaikuttaa urheilijoihin?”

Jatkotutkimus voisikin käsitellä edellä mainittua tutkimuskysymystä, eli sitä, miten urheilija voi hyötyä hänestä itsestään tai hänen pelistään kerätystä datasta. Tällaisella tutkimuksella voisi olla hyvinkin merkityksellisiä tuloksia siitä, miten datan hyödyntäminen vaikuttaa yksittäisen urheilijan suorituksiin, kehittymiseen fyysisesti ja siihen voisiko sillä auttaa pelaajaa myös ymmärtämään peliä paremmin.

Tutkielma osoitti, että NFL käyttää dataa pelaajien turvallisuuden parantamiseen ja loukkaantumisriskien pienentämiseen. Jatkotutkimus pelaajien turvallisuuteen liittyvästä datasta olisi hyvin ajankohtaista. Mikään aikaisempi tutkimus ei todistanut, että esimerkiksi muissa kontaktilajeissa, kuten jääkiekossa käytettäisiin dataa loukkaantumisriskien havaitsemiseen. Toinen laji, missä loukkaantumisriskejä tarkasteltiin pelaajista kerätystä datasta, oli jalkapallo.

Siruteknologian kehittyessä voisi ajatella jatkotutkimuksia myös sen kannalta, mihin kaikkiin urheilijan varusteisiin voisi kiinnittää sensoreita siten, että sensoreista saatu data olisi hyödyllistä. Onko esimerkiksi mahdollista, että jääkiekkokypärään tai amerikkalaisen jalkapallon pelaajan kypärään kiinnitettäisiin sensorit, jotka havaitsisivat päähän kohdistuneita iskuja? Voitaisiinko esimerkiksi kypärissä olevilla sensoreilla saada dataa mahdollisista aivotärähdyksistä tai tutkia voimaa, mikä pelaajan päähän kohdistuu taklaushetkellä. Jatkotutkimuskysymys voisi olla myös: pystytäänkö jo nyt pelistä kerättävällä datalla huomioida taklaavan pelaajan vauhti ja voima osuessa taklattavaan pelaajaan niin, että taklauksesta saataisiin dataa mihin suurin voima on keskittynyt ja onko mahdollisesti taklattava pelaaja altistunut aivotärähdykselle?

Tutkielma osoittaa, että aiheena datan, data-analytiikan ja teknologian hyödyntäminen huippu-urheilussa on ajankohtainen, mutta ei valmis. Dataa osataan kerätä, mutta osataanko sitä vielä hyödyntää riittävällä tasolla? Tuleeko tekoälystä ja koneoppimisesta urheilijoiden seuraava joukkueoveri? Tulevaisuus näyttää, mutta tämä tutkielma esittää, että data-analytiikan ja teknologian hyödyntäminen on välttämätöntä huippu-urheilussa pärjätäkseen. Huippu-urheilussa dataa hyödynnetään jo nyt paljon, eikä se ole tulevaisuudessa pienentymässä, vaan päinvastoin.

Lähteet

Ahonen, S-M., Jääskeläinen, P., Kangasniemi, M., Liikanen, E., Pietilä, A-M. & Utriainen, K. 2013. Kuvaileva kirjallisuuskatsaus: eteneminen tutkimuskysymyksestä jäsennettyyn tietoon. Luettavissa:

[https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_proquest_journals_1469873650&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Ahonen,%20S-](https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_proquest_journals_1469873650&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Ahonen,%20S-M.,%20Jääskeläinen,%20P.,%20Kangasniemi,%20M.,%20Utriainen,%20K.,%20Pietilä,%20A-M.%20%26%20Liikanen%20E.%202013.%20Kuvaileva%20kirjallisuuskatsaus:%20eteneminen%20tutkimuskysymyksestä%20jäsennettyyn%20tietoon.&offset=0)

[M.,%20Jääskeläinen,%20P.,%20Kangasniemi,%20M.,%20Utriainen,%20K.,%20Pietilä,%20A-](https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_proquest_journals_1469873650&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Ahonen,%20S-M.,%20Jääskeläinen,%20P.,%20Kangasniemi,%20M.,%20Utriainen,%20K.,%20Pietilä,%20A-M.%20%26%20Liikanen%20E.%202013.%20Kuvaileva%20kirjallisuuskatsaus:%20eteneminen%20tutkimuskysymyksestä%20jäsennettyyn%20tietoon.&offset=0)

[M.%20%26%20Liikanen%20E.%202013.%20Kuvaileva%20kirjallisuuskatsaus:%20eteneminen%20tutkimuskysymyksestä%20jäsennettyyn%20tietoon.&offset=0](https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_proquest_journals_1469873650&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Ahonen,%20S-M.,%20Jääskeläinen,%20P.,%20Kangasniemi,%20M.,%20Utriainen,%20K.,%20Pietilä,%20A-M.%20%26%20Liikanen%20E.%202013.%20Kuvaileva%20kirjallisuuskatsaus:%20eteneminen%20tutkimuskysymyksestä%20jäsennettyyn%20tietoon.&offset=0) Luettu

11.9.2021. Luettu: 10.9.2021.

Alam, T. & Almulla, J. 2020. Machine Learning Models Reveal Key Performance Metrics of Football Players to Win Matches in Qatar Stars League. Luettavissa:

https://helsinki.primo.exlibrisgroup.com/discovery/fulldisplay?docid=cdi_doaj_primary_oai_doaj_org_article_94e56408ccb14b308961b7f847c1657c&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Machine%20Learning%20Models%20Reveal%20Key%20Performance%20Metrics%20of%20Football%20Players%20to%20Win%20Matches%20in%20Qatar%20Stars%20League. Luettu: 2.10.2021.

Avery, E. 2021. An Introduction to Modeling NFL Tracking Data. Luettavissa:

<https://repository.yu.edu/bitstream/handle/20.500.12202/5659/Avery%20Ennis%20Thesis%20An%20Introduction%20to%20Modeling%20NFL%20Tracking%20Data.pdf?sequence=1&isAllowed=y>. Luettu: 11.10.2021.

Bernard, M. 2016. Big Data in Practice: How 45 successful companies used big data analytics to deliver extraordinary results. John Wiley Sons Inc. New York. Luettavissa:

[https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_askewsholts_vlebooks_9781119231417&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Bernard,%20M.%202016.%20Big%20Data%20in%20Practice:%20How%2045%20successful%20companies%20used%20big%20data%20ana-lytics%20to%20deliver%20extraordinary%20results.&offset=0](https://helka.helsinki.fi/discovery/fulldisplay?docid=cdi_askewsholts_vlebooks_9781119231417&context=PC&vid=358UOH_INST:VU1&lang=fi&search_scope=MyInst_and_CI&adaptor=Primo%20Central&tab=Everything&query=any,contains,Bernard,%20M.%202016.%20Big%20Data%20in%20Practice:%20How%2045%20successful%20companies%20used%20big%20data%20analytics%20to%20deliver%20extraordinary%20results.&offset=0). Luettu: 19.9.2021.

Cai, L. & Zhao, C. 2021. Situation and lessons of application of NBA big data technology. Luettavissa: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9533143>. Luettu: 26.9.2021.

Curro, J. & Schuckers, M. 2013. Total Hockey Rating (tHoR): A comprehensive statistical rating National Hockey League forwards and defensemen based upon all on-ice event. Luettavissa: https://www.statsportsconsulting.com/wp-content/uploads/Schuckers_Curro_MIT_Sloan_THoR.pdf. Luettu: 17.10.2021.

Dehesa, Vaquera, Gonçalves, Mateus, Gomez-Ruano & Sampaio. 2019. Key Game Indicators in NBA Players' Performance Profiles. Luettavissa: <https://hrcak.srce.hr/ojs/index.php/kinesiology/article/view/5456>. Luettu: 1.10.2021.

Ghasem, W., Saxon, L. & Valenzuela, J. 2021. Player Tracking Technology and Data for Injury Prevention in the National Football League. Luettavissa: <https://oce-ovid-com.libproxy.helsinki.fi/article/00149619-202109000-00003/PDF>. Luettu: 17.10.2021.

Goldsberry, K. 2019. Data Science and the 3-point Revolution in the NBA. Luettavissa: <https://ageofrevolutions.com/2019/02/25/data-science-and-the-3-point-revolution-in-the-nba/>. Luettu: 22.10.2021.

Lemmettely, T. 2016. Puettavat sensorijärjestelmät urheilussa. Pro gradu -tutkielma. Itä-Suomen yliopisto, Luonnontieteiden ja metsätieteiden tiedekunta. Luettavissa: https://erepo.uef.fi/bitstream/handle/123456789/17272/urn_nbn_fi_uef-20161313.pdf?sequence=-1. Luettu: 25.9.2021.

Miller, T. 2015. Sports Analytics and Data Science: Winning the Game with Methods and Models. Pearson. Old- Tappan, New Jersey. E-kirja. Luettu: 14.9.2021.

Mincev, S. 2020. Analysing Data Mining Methods in Sports Analytics: A Case Study in NHL Salary Prediction. Luettavissa: <https://run.unl.pt/handle/10362/117866>. Luettu: 15.10.2021.

Mäntynen, N. 2018. IoT – Asioiden yhdistäminen. Pro gradu -tutkielma. Itä-Suomen yliopisto, Luonnontieteiden ja metsätieteiden tiedekunta. Luettavissa: https://erepo.uef.fi/bitstream/handle/123456789/19139/urn_nbn_fi_uef-20180233.pdf?sequence=-1. Luettu: 30.10.2021.

NBA. 2021. Luettavissa: <https://www.nba.com/stats/player/201939/>. Luettu 4.11.2021.

Pan, S., Vasilakos, A., Wang, J. & Zhou, L. 2017. Machine learning on big data: Opportunities and challenges. Luettavissa: <https://www-sciencedirect-com.ezproxy.haaga-helia.fi/science/article/pii/S0925231217300577?via%3Dihub>. Luettu: 23.9.2021.

Parmenter, D. 2015. Key performance indicators: developing, implementing, and using winning KPIs. Wiley. Hoboken New Jersey. E-kirja. Luettu: 18.9.2021.

Randén, E. 2017. Pieni sanakirja tekoälystä. Luettavissa:

<https://www.cgi.com/fi/fi/blogi/pieni-sanakirja-tekoalysta>. Luettu: 25.9.2021.

Runkler, T. 2016. Data Analytics. Models and Algorithms for Intelligent Data Analysis.

Springer Vieweg. Wiesbaden. E-kirja. Luettu: 29.9.2021.

Ryhänen, A. 2020. Data-Analytiikan hyödyntäminen urheilujoukkueen suorituskyvyn kehittämisessä. Kandidaatintyö. Tampereen yliopisto, tekniikan ja luonnontieteiden tiedekunta. Luettavissa:

<https://trepo.tuni.fi/bitstream/handle/10024/122915/RyhänenAntti.pdf?sequence=2>. Luettu: 10.9.2021.

Räikkönen, L. 2021. Puettavat teknologiat liikunnassa ja urheilussa. Kandidaatintyö.

Jyväskylän yliopisto, informaatioteknologian tiedekunta. Luettavissa:

<https://jyx.jyu.fi/bitstream/handle/123456789/77970/URN%3aNBN%3afi%3ajyu-202109305034.pdf?sequence=1&isAllowed=y>. Luettu: 18.10.2021.

Sarlis, V., Tjortjis, C. 2020. Sport analytics – Evaluation of basketball players and team performance. Luettavissa:

<https://www.sciencedirect.com/science/article/abs/pii/S0306437920300557>. Luettu: 20.9.2021.

Swartz, T. 2017. Hockey Analytics. Luettavissa:

<https://www.sfu.ca/~tswartz/papers/statsref.pdf>. Luettu: 4.10.2021.

Valle C. Athlete Tracking – What's New in Sports Technology. Luettavissa:

<https://simplifaster.com/articles/athlete-tracking-sports-technology/>. Luettu: 5.11.2021.

Vuorimies, W. 2018. Urheiluanalytiikan hyödyntäminen koripallossa, Kandidaatintyö.

Tampereen teknillinen yliopisto. Kandidaatintyö. Tampereen yliopisto, teknillinen tiedekunta. Luettavissa:

<https://trepo.tuni.fi/bitstream/handle/123456789/26300/Vuorimies.pdf?sequence=4&isAllowed=y>. Luettu: 9.9.2021.

Wisehockey A game changer: RealTime Sport Analytics Platform. Overview Wisehockey system. 2020. Luettavissa: https://wisehockey.com/wp-content/uploads/2020/02/wh_overview_en.pdf.

Luettu: 6.10.2021.