

Antti Kuusela

**ESINEIDEN INTERNETIN TIEDONHALLINNAN  
OMINAISPIIRTEET**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2017

# TIIVISTELMÄ

Kuusela, Antti

Esineiden internetin tiedonhallinnan ominaispiirteet

Jyväskylä: Jyväskylän yliopisto, 2017, 38 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Taipalus, Toni

Kandidaatintutkielma käsittelee esineiden internetin tiedonhallinnan ominaispiirteitä. Tutkielmassa määritellään esineiden internetin tiedonhallinnan erilaiset kokonaisuudet ja niihin liittyvät haasteet sekä arvioidaan esineiden internetin tiedonhallinnan mahdollistavia sovelluksia ja palveluita. Tiedonhallinnan eri ulottuvuuksia käsitellään tutkielmassa yhtenäisen viitekehyksen perusteella. Esineiden internetin tiedonhallinnan ominaispiirteitä arvioidaan tutkielmassa teknologisten ratkaisujen, havainnollistavien mallien sekä erilaisten kontekstien näkökulmista.

Tutkimuksen tarkoituksena on muodostaa kokonaisvaltainen ja hyödyllinen käsitys esineiden internetin tiedonhallinnasta sekä siihen liittyvistä teknologisista ja liiketoiminnallisista ominaisuuksista. Tutkimuksen tulosten perusteella kehitettiin kattava analyysi esineiden internetin tiedonhallinnasta. Keskeisimpinä tuloksina löydettiin esineiden internetin tiedonhallintaa edistäviä arkkitehtuureja, malleja sekä tietoturvaan ja tiedon integrointiin pohjautuvia ratkaisuja. Esineiden internetin tiedonhallinnan haasteiksi muodostuivat ongelmat tiedonsiirrossa, tietoturvassa, standardeissa ja liiketoiminnallisissa erityispiirteissä. Lisäksi tulosten perusteella arvioitiin pilvialusta- ja datakeskusratkaisujen sopivuutta esineiden internetin tiedonhallinnan tukemiseksi. Kandidaatintutkielma on toteutettu kirjallisuuskatsauksena ja tutkimus perustuu pääosin akateemisten julkaisujen tieteellisiin artikkeleihin.

Asiasanat: esineiden internet, tiedonhallinta, big data, datakeskeisyys, RFID, pilvilaskenta

## ABSTRACT

Kuusela, Antti

The characteristics of data management of internet of things

Jyväskylä: University of Jyväskylä, 2017, 38 p.

Information Systems, Bachelor's Thesis

Supervisor(s): Taipalus, Toni

The Bachelor's Thesis covers the characteristics of data management of internet of things. In the study, the different aspects of data management of internet of things, various challenges based on those aspects and the enabling applications and services are being defined and evaluated. The perspectives of data management are being described by using a unified framework. In the study, the characteristics of data management of internet of things are being estimated by reviewing versatile technological applications, illustrative models and different contexts.

The object of the study is to create a comprehensive and useful understanding about the data management of internet of things and its technological and business features. Based on the results, an extensive analysis of the data management of internet of things was developed. The most essential results were the architectures, models and solutions, developed for data security and data integration that improved the data management of internet of things. Problems with data transfer, data security, standards and business characteristics were defined as the challenges of data management of internet of things. In addition, the applicability of various cloud platform and datacenter solutions of the data management of internet of things was reviewed based on the results. The Bachelor's Thesis has been executed as a literature review and the research is founded mainly on scientific articles of academic journals.

Keywords: internet of things, data management, big data, data-centric, RFID, cloud computing

## **KUVIOT**

KUVIO 1 IoT-sovellusten toimialat .....	8
KUVIO 2 Tiedon hierarkia - DIKW hierarchy .....	10
KUVIO 3 Tiedonhallinta DAMA - DMBOK2 Guide Knowledge Area Wheel ...	11

## **TAULUKOT**

TAULUKKO 1 IoT-järjestelmän SOA-arkkitehtuuri .....	15
TAULUKKO 2 IoT-tiedonhallinnan sovelluksia ja palveluita .....	30

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

KUVIOT

1	JOHDANTO .....	6
1.1	Esineiden internet .....	7
1.2	Tiedonhallinta.....	9
2	ESINEIDEN INTERNETIN TIEDONHALLINTA.....	13
2.1	Tietohallinto ja tietoarkkitehtuuri.....	13
2.2	Tiedon mallinnus, dokumentointi ja tallennus.....	15
2.3	Tietoturva ja protokollat .....	17
2.4	Tiedon integrointi, varastointi ja analysointi.....	18
3	ESINEIDEN INTERNETIN TIEDONHALLINNAN HAASTEET .....	20
3.1	Teknologiset haasteet .....	20
3.1.1	Tietoliikenne.....	21
3.1.2	Tietoturva ja yksityisyys.....	22
3.1.3	Standardit .....	23
3.2	Liiketoiminnalliset ja yleiset haasteet .....	23
4	IOT-TIEDONHALLINNAN SOVELLUKSIA JA PALVELUITA .....	25
4.1	Pilvipalvelut.....	25
4.2	Datakeskukset.....	27
5	YHTEENVETO .....	31
	LÄHTEET.....	34
	KAUPALLISET LÄHTEET .....	38

# 1 JOHDANTO

Teknologisen kehityksen ansiosta ihmiset voivat suorittaa jokapäiväisen arki- ja vapaa-ajan elämänsä toimintoja langattomasti internetin välityksellä. Miljardeit ihmiset ympäri maapalloa hyödyntävät internetin mahdollistamia monipuolisia palveluita niin työssä, opiskelussa kuin sosiaalisessa verkostoitumisessa ja ajanvietossa (Miorandi, Sicari, De Pellegrini & Chlamtac, 2012). Toisaalta jatkuvasti kehittyvät älylaitteet ja älylliset ominaisuudet ihmisten käyttämässä laitteissa lisääntyvät yhä enemmän tulevaisuudessa, mikä mahdollistaa fyysisen ja virtuaalisen maailman yhdistymisen langattomasti liitoksissa olevien yksilöllisten laitteiden ja esineiden kommunikoinnin (Welbourne ym., 2009). Esineiden välistä, autonomisesti ilman ihmiskäyttäjää toimivaa, verkottumista ja tietojen analysointia kutsutaan esineiden internetiksi (engl. Internet of Things, IoT), joka on kattava paradigma kuvaamaan monimuotoista ja alati laajenevaa kokonaisuutta langattomasti ja tavoitteellisesti keskenään kommunikoivista esineistä (Atzori, Iera & Morabito, 2010).

Koska esineiden internetissä toimivien laitteiden ja niiden tuottaman erimuotoisen ja -laatuisen datan määrä kasvaa yhä merkittävämmiin tulevaisuudessa, täytyy IoT-ympäristön tiedonhallinta toteuttaa tehokkaaksi ja kestäväksi ratkaisuksi (Sun & Jara, 2014). Kerätyn ja analysoidun datan tulokset hyödyntäminen on aikaansaanut datakeskeisyyden (engl. data-centric) liiketoiminnan, kotitalouksien ja sosiaalisten palvelujen tuottamisessa ja kehittämisessä, mikä korostuu esineiden internetin laajentuessa ja levitessä vieläkin useammille elämän eri osa-alueille (Aggarwal, Ashish & Sheth, 2013).

Tämän kandidaatintutkielman tavoitteena on määritellä esineiden internetiin erityisesti liittyviä ominaisuuksia ja näkökulmia tiedonhallinnan kannalta sekä integroida IoT-tiedonhallinnan eri ulottuvuudet organisaation kilpailuasemaa ylläpitäviksi ja kehittäviksi tekijöiksi. Tutkimusongelmat kirjallisuuskatsauksena toteutettavassa tutkimuksessa ovat:

- *Mitä ominaispiirteitä esineiden internetin tiedonhallintaan liittyy?*
- *Mitä haasteita liittyy esineiden internetin tiedonhallintaan?*
- *Millaisia sovelluksia ja palveluita on kehitetty esineiden internetin tiedonhallintaa varten?*

## 1.1 Esineiden internet

Esineiden internet (engl. Internet of Things, IoT) on merkitykseltään laaja paradigma, joka yhdistää lukuisat teknologiset ratkaisut ja palvelut yhden termin ja kokonaisuuden alaisuuteen. Vaikka nykyään ihmiset kommunikoivat ja käyttävät internetiä yhä enemmän, kasvaa erilaisten älylaitteiden ja fyysisten esineiden sekä niistä muodostuvien verkostojen välinen tiedonsiirto, analytiikka ja yhteistyö tulevaisuudessa merkittävästi. (Miorandi ym., 2012.)

Kevin Ashton esitteli termin esineiden internet - yhtäläisesti suomenkielessä myös asioiden internet - vuonna 1999, ja hän määritteli IoT-järjestelmien koostuvan radiotaajuista identifiointiteknologiaa (engl. Radio Frequency Identification, RFID) hyödyntävistä yksilöllisistä ja yhteensopivista esineistä (Li, Da Xu, Zhao, 2015). Laitteiden ja esineiden RFID-tagit, eli pienikokoiset merkit, siirtävätkin IoT-järjestelmässä langattomasti dataa RFID-lukijoille, joiden avulla voidaan tehokkaasti analysoida ja luoda hyödyllisiä palveluita niin yrityksille kuin yksityisille käyttäjille. Langattomat sensoriverkot (engl. wireless sensor network, WSN) toimivat myös oleellisena teknologisena kokonaisuutena IoT-ympäristössä, mikä liittyy ja integroi useiden fyysisten esineiden älysensorit yhteentoimivaksi verkoksi, joka tukee esineiden toimintaympäristön tunnistusta ja yhtenäistä valvontaa. (Da Xu, He & Li, 2014.) WSN-ratkaisut sisältävät myös ohjaimia, jotka kohdistavat ja siirtävät tehokkaasti sensoridatan IoT-järjestelmän eri toimijoille (Atzori ym., 2010).

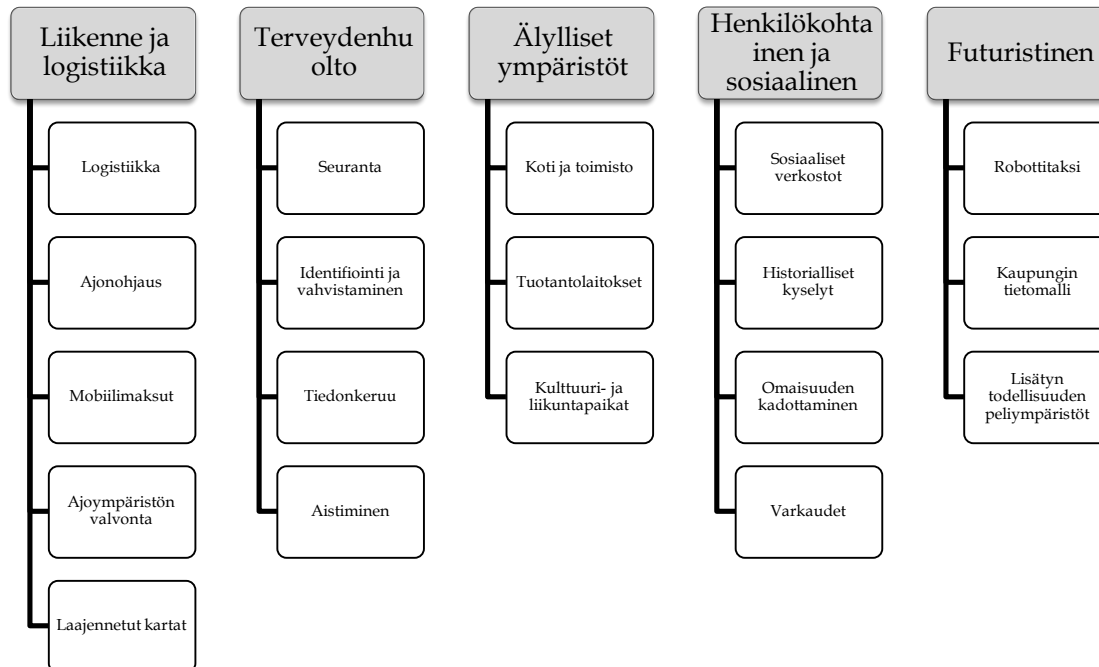
Esineiden internetin katsotaan perustuvan sille merkittävän RFID-teknologian ohella lukuisista muista teknologisista ratkaisuista. Atzori ym. (2010) jaottelevat IoT-paradigman kolmeen keskeiseen näkökulmaan, joihin esineiden internetin toteutukset, sekä siihen liittyvät tutkimukset (Aggarwal ym., 2013), painottuvat eri tavoin: esineisiin, internetiin ja semanttisuuteen.

Esineiden näkökulmasta IoT-ympäristö koostuu sen mahdollistavista teknologioista, kuten fyysisten esineiden RFID-merkeistä, lähitunnistusteknologioista (engl. Near Field Communication, NFC), älylaitteista sekä muista arkisista asioista ja esineistä. Myös sähköisiä tuotekoodeja (engl. Electronic Product Code, EPC) sekä RFID-tageja yhdistävä hajautettu tietojärjestelmä EPCglobal Network on merkittävä osa-alue esineiden internetiä (Fabian & Günther, 2009). Internetin osalta IoT-kokonaisuutta voidaan tarkastella esineiden webin (engl. Web of Things, WoT), matalatehoisten laitteiden IP-osoiteteknologioiden ja -protokollien Internet 0 -ratkaisujen sekä älyllisten esineiden IP-protokollien (engl. IP for Smart Objects, IPSO) kannalta. Semanttisesta eli esineiden erilaisten merkityksien näkökulmasta IoT-kokonaisuutta käsiteltäessä keskitytään semanttisten teknologioiden, datan perustelemisen ja pohdinnan sekä semanttisuuden perustuvien IoT-ympäristöjen toteuttamiseen. Keskeisenä ominaisuutena esineiden internetin mahdollistajina ovat älykkäästi ja semanttisesti toimivat väliohjelmistot sekä esineiden välinen kommunikointi ja liitettävyys. (Atzori ym., 2010.) Yleisesti Gubbi, Buyya, Marusic ja Palaniswami (2013) toteavat IoT-sovellusten integroinnin ja yhteistoiminnan mahdollistajana on pilvilaskenta (engl. Cloud Computing), jonka ansiosta suuri skaalautuvuus, erilaisten alus-

tojen hyödyntäminen, langattomien sensoriverkkojen monipuolinen käyttäminen sekä tietojenkäsittelyn tehokkuus on nopeasti ja luotettavasti toteutettavissa.

Varsinkin IoT-aikakaudella kontekstittietoinen tietojenkäsittely on erittäin tärkeässä tehtävässä kaikkialle leviävän ja kaikkialla läsnä olevan tietotekniikan tukemisessa erilaisissa sulautetuissa järjestelmissä, vaikka teknologiana – kuten monet muut IoT-teknologiat – kontekstittietoisuutta on hyödynnetty jo 1990-luvulta alkaen (Perera, Zaslavsky, Christen & Georgakopoulos, 2014). Esineiden internetin onkin arvioitu laajenevan merkittävästi tulevaisuudessa, sillä Business Insiderin raportin mukaan vuoteen 2020 mennessä internetissä on yhteensä 34 miljardia laitetta, joista esineiden internetiin liittyviä laitteita on 24 miljardia. IoT-ratkaisuihin käytetään raportin mukaan 2016 alkaen seuraavan viiden vuoden ajan 6 biljoonaa dollaria, koska IoT-ratkaisujen koetaan kehittävän kustannustehokkuutta, suorituskykyä, uusien palveluiden luomista, markkinoille pääsyä sekä valtion näkökulmasta kansalaisten elämänlaatua. (Greenough, 2016.)

Esineiden internet sisältää useita eri käyttötarkoituksiin perustuvia sovelluksia ja kokonaisuuksia, jotka osallistuvat hyödyllisten palveluiden kehittämiseen ja tuottamiseen nykyajan yhteiskunnassa. Atzori ym. (2010) määrittelevät IoT-sovellusten toimintaympäristön (KUVIO 1) jakautuvan liikenteeseen ja logistiikkaan, terveydenhuoltoon, älylliseen ympäristöön, henkilökohtaiseen ja sosiaaliseen sekä futuristisiin ratkaisuihin.



KUVIO 1 IoT-sovellusten toimialat (Atzori ym., 2010)

Aggarwal ym. (2013) määrittelevät sovellusten jakautuvan tuotteiden valmistamisen ja logistiikan seurantaan, älylliseen ympäristöön, sosiaalisen tunnistamiseen sosiaalisissa verkostoissa, älylaitteisiin, identifiointiin ja pääsyn valvon-



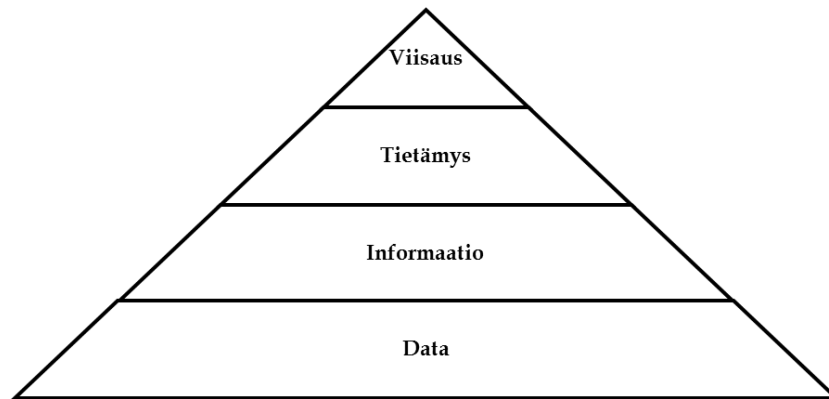
taan, sähköisiin maksujärjestelmiin sekä terveyteen liittyviin sovelluksiin. IoT-ratkaisut mahdollistavatkin useiden erilaisten palveluiden luomisen ja kehittämisen. IoT-sovellukset voivat edistää lentomatkestämisen, autoilun ja liikenteen, tiedonsiirron, terveydenhuollon, vähittäismyynnin, maatalouden, valmistuksen, ympäristön hallinnan, vakuutus- ja rahoituspalveluiden sektoreiden toimintaa monipuolisesti (Bandyopadhyay & Sen, 2011). Esimerkiksi ruoka- ja kaivosteollisuuden turvallisuuden ja paloturvallisuuden (Da Xu ym., 2014) sekä yleisen turvallisuuden ja valvonnan parantamisessa (Miorandi ym., 2012) ja älykaupungin kokonaisuuden luomisessa (Rathore, Ahmad, Paul & Rho, 2016) esineiden internetin laaja-alaiset toteutukset voivat edesauttaa uusien näkökulmien ja tarpeiden tehokasta kehittämistä.

Vaikka IoT-järjestelmien tuottamat mahdolliset hyödyt ovat merkittäviä useista eri näkökulmista, vaatii esineiden internetin tuloksellinen käyttäminen Dijkmanin, Sprenkelsin, Peetersin ja Janssenin (2015) mukaan liiketoiminnallisten ominaisuuksien huomioimista liiketoiminta- ja arvonluontimallien kehittämisessä ja päivittämisessä IoT-aikakaudelle. Lisäksi Ebert, Gallardo, Hernantes ja Serrano (2016) toteavat järjestelmäkehityksen DevOps-käytänteissä, sulautettujen järjestelmien osalta, esineiden internetin ja sen laitteiden välisen kommunikation ja viestinnän (engl. Machine-to-Machine, M2M) tukevan jatkuvaa järjestelmien kehittämistä, toimitusta ja ketteryttä asiakaspalautteen tehokkaamman hyödyntämisen ansiosta. Teknologinen ja liiketoiminnallinen kehitys useilla eri sektoreilla ja toimialoilla korostaakin tutkimuksen tärkeyttä IoT-paradigman kontekstissa.

## 1.2 Tiedonhallinta

Tiedonhallinta käsitteenä jakautuu moniin erilaisiin kokonaisuuksiin, jotka koostuvat useista näkökulmista, tavoitteista ja menetelmistä liittyen tiedon hyödyntämiseen. Tieto suomenkielisenä sanana on suhteellisen laaja, ja usein tieto jaetaan terminä yksityiskohtaisempiin entiteetteihin. Tutkimuskirjallisuudessa tieto on tavallisesti ryhmitelty Ackoffin (1989) määritelmän mukaan nelitasoisesti dataan, informaatioon, tietämykseen ja viisauteen.

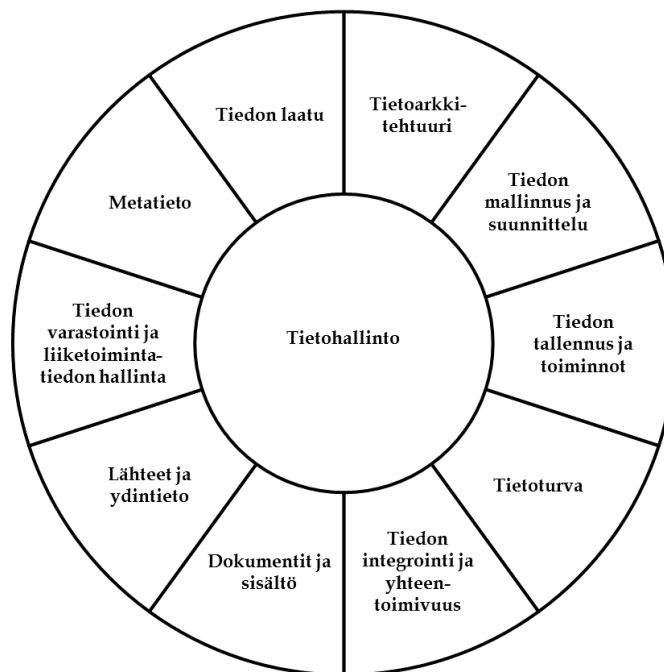
Tiedon määritelmien vaikutussuhteet ja hierarkia voidaan mallintaa pyramidiksi (KUVIO 2), joka havainnollistaa alkuperäisestä – mahdollisesti reaaliaikaisesta kerättävästä – datasta saatavan jonkinlaisen merkityksen omaavaa informaatiota. Analysoimalla rakenteellista ja merkityksellistä informaatiota voidaan lisätä asiantuntemusta sekä varmistaa aiempia oletuksia, minkä avulla on mahdollista toimia tehokkaammin ja viisaammin tekemällä oikeita päätöksiä. (Rowley, 2007.)



KUVIO 2 Tiedon hierarkia - DIKW hierarchy (Rowley, 2007)

Yhtäläisesti tiedonhallinta jaotellaan tutkimuksessa datan, informaation, tietämyksen ja viisauden hallintaan, joita yhdistää suunnitelmallisuus ja tuloksellisuus tiedon keräämisessä, analysoinnissa ja hyödyntämisessä. Tiedonhallintaa tarkasteltaessa datan hallinnan (engl. Data management) näkökulmasta keskitytään organisaation prosesseihin, käytänteisiin, resursseihin ja rakenteisiin, jotka liittyvät tehokkaaseen datan keräämiseen, hyödyntämiseen ja muokkaamiseen tavoitteellisesti (Cupoli, Earley & Henderson, 2014). Informaationhallinta (engl. Information management) puolestaan pyrkii takaamaan kokonaisvaltaisesti käytettävän datan hyödyntämismahdollisuudet liiketoiminnallisten tarpeiden ja vastuiden – kuten suorituskyvyn, asiakkaiden ja prosessien hallinnan tehokkuuden – varmistamiseksi ja kehittämiseksi (Mithas, Ramasubbu & Sambamurthy, 2011). Tietämyksenhallinnan (engl. Knowledge management) näkökulmasta tutkitaan liiketoiminnan eri kokonaisuuksien ja resurssien sidoksuuksia ja ominaisuuksia, jotka vaikuttavat järjestelmällisesti tietämyksen hyödyntämiseen yrityksen liiketoiminnassa (Alavi & Leidner, 2001). Viisautta ja viisaudenhallintaa ei ole käsitetty tietojärjestelmätieteen tutkimuksen kannalta yhtä merkittävänä kokonaisuutena edellä mainittuihin määritelmiin verrattuna, koska viisaudenhallintaan liittyy erittäin abstrakteja, yksilöllisiä, tulkinnanvaraisia sekä vaikeasti mitattavia ja ohjelmoitavia osa-alueita organisaation tiedonhallinnan näkökulmasta (Rowley, 2007).

Esineiden internetin ominaispiirteitä on koneiden välinen kommunikaatio, suurten ja erimuotoisten tietomassojen hallinta, älykkäät ympäristöt sekä tilanetietoinen ja kaikkialla läsnä oleva tietojenkäsittely (Whitmore, Agarwal & Da Xu, 2015). Älylaitteiden kyvykkyydet kerätä, tuottaa ja analysoida dataa IoT-ympäristössä on tiedonhallinnan kannalta merkittävää ja IoT-tiedonhallinnan näkökulmasta kokonaisvaltainen datakeskeisyys vaikuttaa esineiden internetin palvelujen hyödyllisyyteen ja toimivuuteen (Aggarwal ym., 2013). Esineiden internetin datakeskeisyyden ansiosta tässä kandidaatintutkielmassa tarkastellaan esineiden internetin tiedonhallintaa datan hallinnan näkökulmasta, jonka perustana käytetään The Data Management Association (DAMA) -yhdistyksen kehittämää The Data Management Body of Knowledge 2 (DMBOK2) -mallin tietalojen viitekehystä (KUVIO 3, Cupoli, Earley & Henderson, 2014).



KUVIO 3 Tiedonhallinta DAMA - DMBOK2 Guide Knowledge Area Wheel (Cupoli, Earley & Henderson, 2014)

DMBOK2-malli koostuu 11 osa-alueesta, jotka kokoavat yhteen kriittisiä näkökulmia ja kokonaisuuksia organisaation menestyksekkään tiedonhallinnan toteuttamiseksi. Viitekehyksen eri alueet sisältävät prosesseja ja käytäntöjä, jotka täytyy ottaa huomioon tiedonhallinnan ratkaisujen suunnittelussa, käytössä, arvioinnissa ja kehittämisessä. Malli määrittelee myös organisaation tiedonhallintaan liittyviä riippuvuussuhteita ja rooleja, tuloksia ja niiden seurannan menetelmiä sekä resurssien elinkaaren ominaisuuksia. (Cupoli ym., 2014.)

Mallin keskiössä oleva tietohallinto (engl. Data Governance) kokoaa organisaation prosessit, jotka mahdollistavat tiedon ja siihen liittyvien resurssien käytön sekä tiedonhallinnan (Cupoli ym., 2014). Tietohallinto määrittelee organisaation tarpeet ja vastuut tiedonhallintaan keskeisesti liittyvissä päätöksenteon prosesseissa ja käytännöissä (Khatri & Brown, 2010). Tietoarkkitehtuuri (engl. Data Architecture) liittyy yrityksen liiketoiminnallisen ja tiedonhallinnan rakenteellisen arkkitehtuurin yhdeksi kokonaisuudeksi (Cupoli ym., 2014). Tiedon rakenteellinen ja yritysjärjestelmään sidoksinen määrittely yrityksen operatiivisessa tietovarastossa tai tietokannassa onkin merkittävää sekä liiketoiminnan että tiedonhallinnan kannalta (Chalasani & Boppana, 2007). Tiedon mallinnuksella ja suunnittelulla (engl. Data Modeling & Design) tarkoitetaan kaikkia tiedonhallinnan elinkaaren aikaisia prosesseja, jotka Cupoli ym. (2014) määrittelevät tukevan organisaation tietotarpeiden ja -resurssien ”analysointia, suunnittelua, rakentamista, testausta ja ylläpitoa”.

DMBOK2-mallin tiedon tallennus ja toiminnot (engl. Data Storage & Operations) -osa-alue kokoaa organisaation omistamien tietovarantojen mahdollisen varastoinnin ja siihen liittyvien toimintojen hallinnan. Tietoturva (engl. Data Security) määrittelee DMBOK2-rakenteessa kokonaisvaltaisesti eri keinot yk-

sityisen ja luotettavan tietojenkäsittelyn aikaansaamiseksi sekä kuinka luodaan käyttäjille riittävät pääsyoikeudet järjestelmän ja organisaation tietoihin. (Cupoli ym., 2014.)

Tiedon integroinnin ja yhteentoimivuuden (engl. Data Integration & Interoperability) Cupoli ym. (2014) määrittävät tiedonhallinnan DMBOK2-mallissa rakentuvan tiedon "hankinnan, erottelun, muokkauksen, liikuteltavuuden, toimituksen, toisintamisen, liittämisen, virtualisoinnin ja operatiivisen tuen" erilaisiin prosesseihin. Useista eri lähteistä muodostuvan sirpaleisen tietokokonaisuuden hallinta voikin asettaa moninaisia ongelmia tiedonhallinnalle, lisäksi tiedonlaatu ja tiedon jalostaminen, erilaiset tietoskeemat sekä rajoitteet tietolähteissä tekevät tiedon integroinnista laajan kokonaisuuden (Lenzerini, 2002). Nykyään organisaatioilla voi olla käytössään lukuisia järjestelmiä sekä toisistaan poikkeavia laitteistoja ja sovelluksia - jotka kommunikoivat keskenään useissa tietoverkoissa, mikä luo tarpeen semanttiselle verkolle yhteentoimivuuden edistämiseksi (Wong, Ray, Parameswaran & Strassner, 2005).

DMBOK2-mallin dokumentit ja sisältö (engl. Documents & Content) koostaa yhteen tiedonhallinnassa tukena käytettävät dokumentit sekä mahdollistaa rakenteettomien tietolähteiden integroinnin ja toimivuuden rakenteellisten tietokantojen kanssa (Cupoli ym., 2014). Lähteet ja ydintieto (engl. Reference & Master Data) pyrkii vähentämään päällekkäisyyttä sekä korostamaan laatua tiedonhallinnassa standardoiduilla määrittelyillä ja tietoarvoilla (Cupoli ym., 2014). Tiedon varastointi ja liiketoimintatiedon hallinta (engl. Data Warehousing & Business Intelligence) edesauttaa organisaation raportointia ja analysointia hallinnoimalla päätöksenteossa käytettäviä tietoja ja analyttistä prosessointia (Cupoli ym., 2014). Liiketoimintatiedon hallinta ja tietovarastojen hyödyntäminen liiketoiminnassa on korostunut nykyaikana, sillä big datan ja esineiden internetin kehittymisen vaikutuksesta yritysjohtajat haluavat ymmärtää markkinoitaan ja asiakkaitaan yhä kattavammin tekemällä tehokkaita päätöksiä oikeaan aikaan (Chen, Chiang & Storey, 2012).

Metatiedon (engl. Meta-data) kokonaisuus käsittää Cupolin ym. (2014) mukaan organisaation metatietojen "keräämistä, kategoriointia, ylläpitoa, integrointia, kontrollia, hallintaa ja jakamista". Tiedon laatu (engl. Data Quality) pyrkii DMBOK2-mallissa varmistamaan käytettävien tietojen oikeellisuuden valvomalla ja ylläpitämällä laadunhallinnan määritelmiä ja prosesseja sekä kehittämällä tiedonlaatuun liittyviä vaatimuksia. (Cupoli ym., 2014.)

## 2 ESINEIDEN INTERNETIN TIEDONHALLINTA

Esineiden internet muodostuu monipuolisista teknologisista ratkaisuista, mal- leista, protokollista ja käytänteistä, mikä asettaa tarpeen IoT:n kokonaisvaltai- selle tarkastelulle (Li ym., 2015). Miorandi ym. (2012) määrittävät esineiden in- ternetin järjestelmien tärkeimmiksi ominaispiirteiksi ja osa-alueiksi heterogee- niset laitteet, skaalautuvuuden, kaikkialla läsnä olevan kommunikoinnin, ener- giatehokkuuden, esineiden yksilöitävyyden ja niiden kyvyn organisoitua itses- tään, semanttisen yhteentoimivuuden ja datan hallittavuuden sekä järjestel- mään sulautetun yksityisen ja tietoturvallisen tietojenkäsittelyn. Tässä luvussa käsitellään esineiden internetin tiedonhallinnan ominaisuuksia DMBOK2- viitekehyksen tietämysalueiden mukaisesti (kts 1.2 Tiedonhallinta).

### 2.1 Tietohallinto ja tietoarkkitehtuuri

Tietohallinnon päätöksenteon kannalta esineiden internet luo potentiaalisen, sekä määrällisesti että laadullisesti merkittävien tietojen perusteella tehtävän, organisaation strategisen johtamisen (Zimmermann, Schmidt, Sandkuhl, Wißotzki, Jugel & Möhring, 2015). Koska esineiden internetin laitteet voivat toimia laaja-alaisesti eri toimialoilla riippuen organisaation liiketoiminnasta, pystytään IoT-ympäristön keräämään, analysoimaan ja tuottaman datan perus- teella tekemään yksilöllisiä ja reaaliaikaisia päätöksiä sekä kehittämään toimin- tatapoja ja palveluita liittyen organisaation ja IoT:n toimintaan (Aggarwal ym., 2013). Organisaatioiden tietohallinnoilla voi olla myös erilaisia tavoitteita ja odotuksia IoT:n käyttöönottoa kohtaan. Päätöksentekoon vaikuttaa sekä strate- gian reaktiivinen tai proaktiivinen suuntautuminen että liiketoiminnassa mark- kinoiden vetovoimaisuus ja teknologian luomat mahdollisuudet (Li, Hou, Liu & Liu, 2012).

Esineiden internetin tietoarkkitehtuurille on esitetty useita ratkaisuja, jotka rakenteellistavat monimuotoista IoT-infrastruktuuria ja -kokonaisuutta toisis- taan poikkeavilla toimialoilla eri tavoin. Tietoarkkitehtuuriin vaikuttaa myös näkökulma, josta esineiden internetiä tarkastellaan. Da Xu ym. (2014) määritte-

levät nelikerroksisen IoT-palveluorientoituneen arkkitehtuurin (engl. service-oriented architecture, SOA) muodostuvan tunnistus-, verkko-, palvelu- ja rajapintakerroksesta (TAULUKKO 1). Akhbar, Chang, Yao ja Méndez Muñoz (2016) ehdottavat lähes samankaltaista ratkaisua IoT-arkkitehtuurille, joka muodostuisi tunnistus-, verkko-, alusta- ja sovelluskerroksesta. Erilaiset IoT-ratkaisut aikaansaavat myös monimuotoisia tietoarkkitehtuureja. EPCglobal- ja GS1 -organisaatioiden kehittämä sähköinen tuotekoodi (EPC) perustuisi kokonaisuuteen, jossa tuotteet käyttäisivät RFID-tageja ja jokaisella tuotteella olisi yksilöllinen EPC-koodi, jota hyödyntämällä voisi luoda erilaisia palveluita käyttäjille (Weber, 2010). Toisaalta RFID-merkit ovat ainoastaan yksi osa laajemmasta IoT-teknologiainfrastruktuurista (Atzori ym., 2010). Pilvipalvelu Cloud4sens muodostuisi valvottavista infrastruktuureista, kuten älypuhelimista, älysensoreista ja tietokannoista, jotka yhdistetään pilvipalveluun, mistä Cloud4sens-käyttäjät voivat tarkastella tietoja web- ja työpöytäsovellusten avulla (Fazio & Puliafito, 2015). IoT:n sumulaskennassa (engl. Fog Computing) arkkitehtuuri rakentuu nelikerroksisesti älylaitteisiin, langattomiin verkkoihin, tiedonsiirtoprotokolliin ja -menetelmiin sekä datakeskuksiin ja pilvipalveluihin, joissa IoT-analytiikka toteutetaan hyödyntämällä verkonhallinnan menetelmiä ja sovelluksia (Bonomi, Milito, Zhu & Addepalli, 2012).

IoT:n SOA-arkkitehtuurin mukaisesti fyysisiä laitteita lähinnä oleva tunnistuskerros (engl. sensing layer) kerää dataa ja yhdistää laitteita keskenään IoT-teknologioiden - kuten RFID-tagien ja -lukijoiden, sensoreiden, langattomien sensoriverkkojen ja Bluetoothin - sekä protokollien avulla (Da Xu ym., 2014). Tunnistuskerros muodostuu lukuisista laitteista ja esineistä, ja kerroksen kehittämisessä täytyy huomioida datan heterogeenisuus, vaadittavien resursien ja energian käyttäminen, kustannustehokkuus, esineiden kommunikoinnin ja käyttöönoton toteutus sekä millä tavoin esineiden verkosto muodostuu (Li ym., 2015). Verkkokerros rakentuu IoT-palveluiden käyttämistä langallisista ja langattomista tiedonsiirtoyhteyksistä sekä esineitä yhdistävistä infrastruktuureista, kuten internetistä, langattomasta lähiverkosta WLAN:sta, mobiileista- ja sosiaalisista verkoista, langattomista sensoriverkoista sekä tietokannoista. SOA-arkkitehtuurin palvelukerros koostuu palvelun muodostamisesta esineiden tuottaman datan ja liiketoimintalogiikan avulla, palveluiden integroinnista ja jaottelusta yhdistävään palveluväylään sekä palveluiden kokoamisesta yhteiseen tietokantaan, missä väliohjelmistoteknologiat mahdollistavat joustavan ja tehokkaan alustan palveluille. Arkkitehtuurissa ylimpänä olevassa rajapintakerroksessa palveluväylää käytetään käyttöliittymien ja sovellusten ohjelmointirajapintojen kautta, minkä avulla käyttäjät ja sovellukset voivat hyödyntää IoT-palveluita monipuolisesti. Rajapintakerroksessa voidaan käyttää rajapinta-profiileja (engl. interface profile, IFP) sovellusten ja palveluiden määrittelyssä sekä Universal Plug and Play (UPnP) -protokollaa useiden erilaisten esineiden ja sovellusten kommunikoinnissa (Li ym., 2015). Kokonaisuudessaan nelikerroksinen palveluorientoitunut arkkitehtuuri edistää esineiden internetin tietoturvaa erottelemalla toisistaan toiminnot, palvelut, verkot ja laitteet kerroksittain ja keskenään sekä rajoittamalla niiden toiminnan kokonaisvaikutuksia IoT-ympäristössä. (Da Xu ym., 2014.)

TAULUKKO 1 IoT-järjestelmän SOA-arkkitehtuuri (Da Xu ym., 2014)

Kerros	Toiminta	Sisältö
Tunnistuserros	Fyysisen maailman tunnistaminen, valvonta ja datankeruu yhdistämällä IoT-laitteita ja -teknologioita	RFID-tagit, RFID-lukijat, älysensorit, matalatehoiset Bluetooth-laitteet, langattomat sensoriverkot, datan tunnistamisen ja keruun protokollat, muut IoT-teknologiat
Verkkokerros	Verkottumisen ja tiedonsiirron mahdollistaminen langattoman tai langallisen verkon välityksellä	Internet, langattomat sensoriverkot, mobiiliverkot, tietokannat, sosiaalinen verkko, WLAN, pilvipalvelut
Palveluserros	Palveluiden luominen ja hallinta asiakkaiden tarpeiden ja vaatimusten mukaisesti	Liiketoimintalogiikka, palveluiden jaottelu, integrointi ja säilyttäminen, palveluväylä
Rajapintakerros	Käyttäjien ja sovellusten välisen kommunikoinnin tukeminen	Palveluväylä, käyttöliittymä, sovelluksen ohjelmointirajapinta, palvelusopimus, rajapinta

## 2.2 Tiedon mallinnus, dokumentointi ja tallennus

IoT-tiedonhallintaa tarkasteltaessa tiedon mallinnuksen ja suunnittelun sekä dokumentoinnin näkökulmasta arvioidaan perinteisten tietomallien soveltuvuutta esineiden internetin aikaansaamien tiedonhallinnan vaatimusten kanssa. Tiedon mallintamisessa perinteisesti hyödynnetyt mallit eivät käytännöllisyydestänsä huolimatta pysty vastaamaan kaikkiin IoT:n luomiin tarpeisiin. Lisäksi suurten tietomäärien nopeaan käsittelyyn kehitetyt NoSQL-tietokantaperheen avain-arvo-pari-, dokumentti-, sarake- ja graafitietomallit, eivät perinteisten mallien ohella sovellu monipuolisesti esineiden internetin järjestelmien yksityiskohtaiseen tietojenkäsittelyyn. (Sun & Jara, 2014.) Laajalti käytössä oleva relaatiomalli perustuu mukaan kokonaisvaltaiseen ja keskitettyyn taulukoiden hallintaan, mikä voi Leavittin (2010) mukaan IoT:n ja big datan aikakaudella tuottaa ongelmia, kun käsiteltävä data ei ole rakenteellisessa muodossa, tietokannat ovat erittäin hajautettuina osina, tietoa ei voi muokata taulukkomuotoon ja SQL-kyselykieli ei toimi riittävän joustavasti ei-rakenteellisen tiedon kanssa. NoSQL-tietokannat voivat prosessoida nopeasti paljon dataa, mutta SQL:n kaltaisen kyselykielen luominen haastavia prosesseja varten, ACID-ominaisuuksien eli datan atomisuuden, eheyden, eristyneisyyden sekä pysyvyyden varmistaminen, sekä NoSQL-ympäristön osittainen sovellusten ja menetelmien puutteellisuus asettaa haasteita NoSQL-tietokannoille (Leavitt, 2010).

Relaatio- ja NoSQL -mallien ominaisuuksien pohjalta kehitettävä ratkaisukokonaisuus NewSQL pyrkii vastaamaan skaalautuvuuden ja toimintatehokkuuden vaatimusten kasvuun big datan käytön yleistyessä relaatiomallin mukaisilla taulukoilla ja kyselykielellä, mutta NoSQL-mallin kaltaisia hajautettuja pilvilaskennan ominaisuuksia hyödyntäen (Grolinger, Higashino, Tiwari & Capretz, 2013). Tietokantaratkaisujen ja -alustojen, kuten Hadoopin, Apachen Sparkin ja Cassandraan hyödyntäminen on ollut monipuolista IoT-järjestelmien luomisessa, sillä sovellukset mahdollistavat suurien datamäärien prosessoinnin ja analysoinnin tehokkaasti (Díaz, Martín & Rubio, 2016). Relatiomallin mukainen tietokantojen datan kattava normalisointi ei ole NoSQL-tietokannoissa, kuten Cassandra, HBASEssa ja MongoDB:ssä, keskeisintä ja datan toistuvuus, puuttuminen sekä tietomallin epätäydellinen noudattaminen ei vaikuta uusimpien kehitettyjen NoSQL-tietokantojen tehokkaaseen ja hajautettuun toimintakykyyn (Mesiti & Valtolina, 2014).

IoT-tiedonhallinnan kannalta Sun ja Jara (2014) tuovat esille kuusi merkittävää vaatimusta tietomallille: semanttisuuden mahdollistaminen, datan aktiivinen suodattaminen ja analysointi, tietomallin joustavuus ja laajennettavuus, erilaatuisten ja suurten datamassojen hallinta, tiedon yhtenäisen organisoinnin mahdollistaminen tiedonhallintaa varten sekä loogisen ja matemaattisesti oikeellisen kokonaisuuden luominen. Sun ja Jara (2014) määrittelevät IoT-tietomallin rakentuvan kaksitasoisesti kohde ja tapahtuma -kerrokseen, joissa luodaan yksilölliset skeemat ja säännöt erilaisille esineiden internetin kohteille ja tapahtumille. Kohteiden skeemat kuvaillaan tyyppien, muihin kohteisiin verrattavien merkityksien ja toimintojen avulla. Säännöt määrittävät kohteiden merkityssuhteiden, toimintojen ja ominaisuuksien, kuten nimen, toimialueen ja vaatimusten perusteella. Tapahtumien osalta skeemat luodaan myös tapahtumien tyyppien ja muihin tapahtumien merkityksien vertailun kautta sekä säännöt yksilöidään tapahtumien ominaispiirteiden ja merkityssuhteiden avulla. Semanttinen tietomalli sisältää kohteiden suhteet yhdistävän semanttisen linkkiverkon mallin ja tapahtumien suhteellisuuden yhdistävän tapahtumien semanttisen linkkiverkon mallin, jotka rakentuvat esineiden internetin kohteiden ja toiminnan suhteiden loogisten ja matemaattisten yhteyksien kattavaan perusteluun. (Sun & Jara, 2014.)

Esineiden internetin tiedon tallennuksen ja toimintojen määrittelyyn liittyy merkittävästi tietomallien perusteella luotujen tietokantojen rakenteellinen toteutus. IoT-järjestelmiä varten relaatiomallin mukaiset tietokantojen hallintajärjestelmät vaativat paljon muokkausta ja täysin uusien tallennusjärjestelmien luominen IoT-ympäristöön relaatiomallin ohella olisikin tarpeellista, sillä IoT-laitteet voivat toimia myös alueellisesti laajalti ja hajautetusti (Qin, Sheng, Falkner, Dustdar, Wang & Vasilakos, 2016). Suuren hajautetun tallennusjärjestelmän kehittämisessä täytyy huomioida CAP-teoreeman rajoitteet. Hajautettu järjestelmä ei voi samanaikaisesti toteuttaa kuin kahta seuraavista ominaisuuksista: eheys (engl. consistency), saatavuus (engl. availability) ja osituksen sietokyky (engl. partition tolerance) liittyen järjestelmän tilaan ja käsiteltävän datan laatuun. (Brewer, 2000.) Esineiden internetin laajentuessa koneiden välinen kommunikaatio ja tekoäly tiedonhallinnan toimintojen algoritmien, pilvitalennuksen ja -alustojen sekä tallennusinfrastruktuurin analytiikan käyttöönotossa ovat



merkittäviä kokonaisuuksia tiedonhallinnan osa-alueiden joustavuuden ja tehokkuuden kehittämisessä (Gubbi ym., 2013).

## 2.3 Tietoturva ja protokollat

Ensimmäisten IoT-järjestelmien käyttöönotossa tietoturvaan ja tiedon suojaukseen liittyviä ongelmia on ratkaistu tilapäisillä ja tilanneriippuvaisilla sovelluksilla yhden palveluntarjoajan alaisuudessa, mutta tulevaisuuden avoimen IoT-ympäristön ja useiden eri toimijoiden yhteistyö vaatii tietoturvan laajaa hallintaa (Miorandi ym., 2012). IoT-laitteet toimivat säännöllisesti langattomissa verkoissa, joiden salausmenetelmät ovat kehittyneet vuosien varrella, mutta IoT-ympäristön esineiden tehokkuus ei ole tarpeeksi riittävän varman salauksen mahdollistamiseksi (Whitmore ym., 2015). Langattomissa sensoriverkoissa käytettyjä teknologioita voidaan soveltaa IoT:n tietoturvan edistämiseksi (Li ym., 2015), mutta IoT:n joustavan toiminnan kannalta salausalgoritmien täytyy kehittyä sekä nopeudeltaan että energian käytöltään tehokkaammiksi ja salausavaimien jakoa varten täytyy luoda IoT-ympäristöön soveltuva ratkaisu (Bandyopadhyay & Sen, 2011). Organisaation liiketoiminnan näkökulmasta tietoturvaratkaisun täytyy sietää hyökkäyksiä ja toimintahäiriöitä, varmistaa datan oikeellisuus, mahdollistaa tietoihin pääsyn valvonta sekä tukea käyttäjien yksityisyyttä (Weber, 2010).

IoT-järjestelmän toiminta perustuu tehokkaisiin kommunikointiprotokolliin. IoT-ympäristön fyysisten laitteiden ja niiden MAC-osoitteiden kokonaisuus hyödyntää IEEE 802.15.4 -protokollaa, joka tukee 6LoWPAN-protokollan (engl. IPv6 over Low power Wireless Personal Area Networks) datapakettien muokkausta ja sopeutusta verkkokerroksen IPv6-protokollaa varten. Verkkokerroksen reititysprotokolla RPL (engl. Routing Protocol for Low-power and Lossy Networks) mahdollistaa reitityksen muokkauksen IoT-sovelluksien tarpeiden mukaisesti, jolloin sovelluskerroksen CoAP-protokolla (engl. Constrained Application Protocol) tukee useiden sovellusten yhteistoimintakykyä ja kommunikointia. IEEE 802.15.4 -protokollaan kuuluu erilaisia tietoturva-asetuksia riippuen siirrettävän tiedon salauksesta ja datan oikeellisuuden varmistusmenetelmästä. 6LoWPAN-protokollan muokkaustekniikoita varten ei ole kehitetty vielä tietoturvamenetelmiä langattomien IoT-alustojen rajoitusten takia, mutta IPSecin sekä VPN-yhteyksien hyödyntämistä varten on kehitetty IETF-organisaation RFC-dokumentteihin erilaisia pohdintoja ja näkökulmia. RPL-reitityksen on mahdollista toteuttaa joko ilman viestien suojausta, laitteeseen alun perin tehdyllä suojauksella tai verkkoon liittymisen kautta avaimien haltijalta saadulla salatulla avaimella, jolloin laite voi toimia verkon reitittimenä. CoAP-sovellusprotokollassa voidaan käyttää turvallista DTLS-protokollaa salaamalla viestit laitteissa ilman salausta, jaetuilla avaimilla julkisen avaimen avulla lasketulla salausavaimella tai sertifioidulla avaimella. (Granjal, Monteiro & Sa Silva, 2015.)

## 2.4 Tiedon integrointi, varastointi ja analysointi

Tiedon sisällöllinen integrointi ja yhteentoimivuus esineiden internetin kannalta on yksi merkittävimmistä osa-alueista, jotta IoT-järjestelmät olisivat mahdollisia (Atzori ym., 2010). Edellä esiteltyjen arkkitehtuurien, tietomallien, tallennuksen ja tietoturvan lisäksi IoT-ympäristön teknisten ratkaisujen ominaisuudet vaikuttavat laitteiden yhteiseen toimintaan. Esineiden internetissä langattomat sensoriverkot yhdistävät kaikkialla läsnä olevat laitteet verkostoksi, jossa laitteiden tietojenkäsittely, tallennus, visualisointi ja analysointi toteutetaan pilvipalveluna, joka mahdollistaa käyttäjille monipuoliset sovellukset ja toisaalta yhdistää fyysisen ja virtuaalisen maailman (Gubbi ym., 2013). Riippuu paljolti IoT-järjestelmän rakenteesta – millä tavoin järjestelmässä sensorit, passiiviset ja aktiiviset RFID-tagit, RFID-lukijat, sensoriverkon solmut ja laitteet on yhdistetty – kuinka pitkäaikaisesti, laaja-alaisesti ja oikeita tuloksia tuottavasti IoT-järjestelmä voi toimia (Liu, Bolic, Nayak & Stojmenovic, 2008). IoT-järjestelmän tiedon integrointiin liittyy datan puhdistaminen ja muokkaus keskeisesti, sillä RFID-tagien tuottama tieto ei ole usein tarpeeksi monipuolista ja tarkkaa, lisäksi tagien kantaman ja mahdollisten akkujen keston rajallisuus täytyy huomioida kokonaisuuden toimivuuden kannalta (Aggarwal ym., 2013).

Esineiden internetin eri osa-alueiden monipuolinen standardointi vaikuttaa merkittävästi IoT-ympäristön kokonaisvaltaiseen toimivuuteen ja käytönottoon positiivisesti (Li ym., 2015). IoT-tiedonhallinnan näkökulmasta standardointi vaikuttaa lähteiden, tiedon laadun ja metatietojen hallintaan. Standardit tietomallit ja formaatit tukevat tehokasta metatiedon keruuta ja tiedonhallintaa merkittävästi, sillä esineiden internetin laajentuessa tarkat määritelmät ja toimintatavat mahdollistavat automaattisen tietojenkäsittelyn (Miorandi ym., 2012). IoT:n kaikkialla läsnä olevassa ja leviävässä tietotekniikassa keskeistä on kontekstietoisuus, joten tietystä kontekstista kerättävän datan voimassaolon, tarkkuuden ja ajankohtaisuuden määrittäminen – riippuen sensoreiden, datan ja tiedonsiirron laadusta – vaikuttaa tiedon laatuun huomattavasti (Bellavista, Corradi, Fanelli & Foschini, 2012, Pereran ym., 2014 mukaan).

Big data-analytiikka ja suurten tietomassojen hallinta liittyy keskeisenä ominaisuutena IoT-järjestelmän tietojen varastointiin ja liiketoimintatiedon hallintaan. Big data -teknologioiden käyttö IoT-ympäristössä on merkittävää tulevaisuuden kannalta, sillä tiedon prosessointiteho esineiden internetissä ei riitä kerättävää tietomäärää varten, mikä lisää tarvetta sekä big data- että IoT -sovellusten monipuoliselle kehittymiselle. Datakeskuksilla on niin tiedon varastoinnissa kuin keräämisessä ja hallinnassa tärkeä merkitys, sillä datakeskusten täytyy säilyttää ja analysoida turvallisesti yhä enemmän tietoa sekä tuottaa liiketoimintaa kehittäviä hyödyllisiä palveluita ja ratkaisuja perustuen big dataan. (Chen, Mao & Liu, 2014.)

Älykaupungin kontekstissa Rathore ym. (2016) määrittää nelitasoisen IoT big data-analytiikka arkkitehtuurin, joka koostuu tiedon luonnista ja keräämisestä, kommunikoinnista, tiedon hallinnasta ja prosessoinnista sekä tiedon tulokinnasta. Mallin ensimmäisellä tasolla luodaan IoT-sovellusten - kuten älykotiin ja valvontalaitteiden - avulla dataa, joka siirtyy toisen tason verkkoteknolo-

gioiden avulla kolmannen tason tietovarastoinnin ja tiedonhallinnan Hadoop-järjestelmään, jossa tietoja muokataan ja analysoidaan neljännen tason yksityisten ja julkisten käyttäjien sovelluksia varten (Rathore ym., 2016). Nykyään ihmiset käyttävät älypuhelimia monipuolisesti, säännöllisesti ja sosiaalisesti, joten big datan, liiketoimintatiedon hallinnan ja esineiden internetin integrointi voi mahdollistaa myös sosiaalisesti luodun liiketoimintatiedon hallinnan (Chianese, Marulli, Piccialli, Benedusi & Jung, 2017).

### **3 ESINEIDEN INTERNETIN TIEDONHALLINNAN HAASTEET**

Esineiden internetin tiedonhallintaan liittyy useita erilaatuisia ongelmia, jotka vaikuttavat sekä teknisesti että sosiaalisesti IoT-järjestelmään (Whitmore ym., 2015). Jotta esineiden internetin toimintaympäristö voisi kehittyä tehokkaaksi kokonaisuudeksi, täytyy huomioida ja ratkaista tietoverkkoon, tietoturvaan ja yksityisyyteen sekä IoT-laitteiden toimintakyvyn hallintaan liittyviä merkittäviä teknologisia ja liiketoiminnallisia haasteita (Bandyopadhyay & Sen, 2011). Tässä luvussa tarkastellaan esineiden internetin tiedonhallinnan haasteita teknologisista sekä organisaatioon liiketoimintaan vaikuttavista näkökulmista.

#### **3.1 Teknologiset haasteet**

Esineiden internetin teknologinen järjestelmä voi mahdollistaa sulautetun ja kontekstittietoisen tietotekniikan hyödyntämisen monipuolisesti, mutta IoT-ympäristö rakentuu heterogeenisiin laitteisiin ja verkostoihin, joiden ominaisuuksien hallinta luo haasteita IoT-ympäristön teknologisille ratkaisuille (Whitmore ym., 2015). Esineiden internetin kehittämisessä täytyykin huomioida tällä hetkellä käytössä olevien ICT-teknologioiden ja -laitteiden integrointi IoT-järjestelmään (Li ym., 2015). Kaikkialle leviävien IoT-sovellusten verkoston yhteistoiminnan varmistamiseksi täytyy luoda selkeät toteutukset massiivisten datamäärien kustannustehokkaaseen siirtoon, erilaisten resurssien hinnoitteluun ja ihmisten sosiaaliseen yksityisyyteen (Sheng ym., 2013).

Tulevaisuudessa tuotettavan ja käsiteltävän datan määrän huomattava kasvu IoT-palveluiden edellyttämiseksi synnyttää merkittäviä haasteita kehitettävälle alustoille, rakenteelle ja arkkitehtuurille (Akhbar ym., 2016). Tällöin esineiden internet yhdistää miljardeja laitteita keskenään, mikä vaatii niin määrällisesti kuin laadullisestikin joustavaa ja nopeaa tietojenkäsittelyä, joka toteutetaan yhä enemmän pilvipalveluiden avulla (Gubbi ym., 2013). Suurten datamäärien tehokas kontrolli ja analysointi hyödyllisiä palveluita varten korostaa heterogeenisten kokonaisuuksien hallintaa IoT-järjestelmissä (Bandyopadhyay

& Sen, 2011). Heterogeeninen, suuri ja hajautettu datamäärä täytyy huomioida myös tietokantojen suunnittelussa pilvialustoihin, sillä perinteiset relaatiomalliin perustuvat tietokannat eivät kykene ratkaisemaan esineiden internetin aikaansaamia palvelun skaalautuvuuden, saatavuuden, virheiden sietokyvyn ja joustavuuden ongelmia (Grolinger ym., 2013).

IoT-järjestelmän mahdollistavaa sensoridataa varten olisi tarpeellista kehittää monikäyttöiset sensoridatan tunnistus- ja prosessointimekanismit, kommunikointiprotokollat sekä tekniikat tiedonlouhintaa ja analysointia varten (Bandyopadhyay & Sen, 2011). Datan keruussa aktiivisten RFID-tagien akunkeston ja kadonneen datan ongelmat sekä passiivisten RFID-tagien heikon kantaman aiheuttamat virheelliset tai puuttuvat tulokset tekevät sensoridatan hallinnasta ja hyödyntämisestä haasteellista (Aggarwal ym., 2013).

### 3.1.1 Tietoliikenne

Internetin ja tiedonsiirron täytyy kehittyä IoT-ympäristön tarpeisiin soveltuviksi, sillä esineiden internet vaatii yhä kokonaisvaltaisempaa skaalautuvuutta, mobiiliutta sekä verkon saatavuutta ja hallintaa (Bandyopadhyay & Sen, 2011). Esineiden internetiin kytketyt laitteet muodostavat laajoja kokonaisuuksia, mikä asettaa tiedonsiirron kapasiteetille ja toimintakyvylle uudenlaisia vaatimuksia, eikä nykyisten protokollien yhdistämisen ja ruuhkanhallinnan menetelmät ole riittäviä IoT-järjestelmiä varten (Atzori ym., 2010).

Sensoreiden yhdistäminen verkkoon esimerkiksi 3G:n ja WiFin avulla voi olla vaikeaa, koska RFID-tagien akunkesto tai tehokkuus ei riitä verkon jatkuvaan ylläpitoon (Aggarwal ym., 2013). Toisaalta Bluetooth-teknologian versio 5 saattaa moninkertaisilla parannuksillaan kommunikaation nopeuteen, toimintasäteeseen sekä datamäärään viesteissä kehittää IoT-ympäristön tiedonsiirtoa (Bluetooth, 2016). IEEE 802.15.4 -protokollan rajallinen kapasiteetti heikentää IoT-järjestelmän skaalautuvuutta ja sovellusten ruuhkanhallintaa, lisäksi 6LoWPAN-protokollan sisällön tiivistäminen IPv6-paketeissa ei riitä tehostamaan IoT-verkkoa, jos laitteita on verkossa useita miljardeja (Sheng ym., 2013). Verkkokerroksen RPL- ja sovelluskerroksen CoAP -protokollat eivät kestä todella suuria asiakasmääriä, mikä voi asettaa IoT-verkon alttiiksi hajautetuille palvelunestohyökkäyksille (engl. Distributed Denial of Service, DDoS) (Sheng ym., 2013). RFID-tageihin perustuva erittäin suuren määrän tuotteita yhdistävä EPC-arkkitehtuuri voisi myös aiheuttaa DDoS-hyökkäyksiä, koska EPC:ssä käytettävä ONS-protokolla (engl. Object Name Service) hyödyntää DDoS-hyökkäyksissä käytettyä DNS-protokollan haavoittuvuuksia (Fabian & Günther, 2009).

Sensoreiden käyttöikää optimoidaan muokkaamalla niiden tehtäväkiertoa (engl. duty-cycle) ja energiankeruuta, jolloin sensorit toimivat vain tarvittaessa ja kommunikaatio laitteiden välillä keskeytyy. Jotta protokollien pakettien häviämisen ja synkronoinnin hallinta olisi mahdollista, täytyy luoda selkeät standardit tehtäväkierron ja protokollien yhteistoiminnan edistämiseksi. Tulevaisuuden IoT-laitteiden tiedonsiirtotarpeet ja palveluiden laadun vaatimukset voivat myös lisätä ongelmia esineiden internetin protokollien käytössä. Esineiden internetin laitteet vaativat useita erilaisia verkko- ja resurssikokonaisuuksia

tietoliikenteeltä ja protokollien hyödyntämisen kustannukset ja energian kulutus vaikeuttavat IoT-järjestelmän tiedonsiirron toteutusta. (Sheng ym., 2013.) Esineiden internetin tiedonsiirron tehokas toiminta edellyttää uudenlaisten teknologisten ratkaisujen kehittämistä ja tutkimista Kokkonien (2017, 22) kehittämä nanoskaalan sensoriverkkojen malli, joka perustuu 10 000 GHz:n kaistanleveyteen sekä THz-taajuusalueen hyödyntämiseen tiedonsiirrosta voi olla merkittävä kokonaisratkaisu IoT-sovellusten kommunikoinnin ja toiminnan edistäjänä tulevaisuudessa.

### 3.1.2 Tietoturva ja yksityisyys

IoT-laitteiden laskentatehokkuuden ja vähäisten tietoteknisten ominaisuuksien vuoksi aiheutuu tietoturvaan ja käyttäjien yksityisyyteen liittyviä ongelmia. Weberin (2010) mukaan tietoturvan takaamiseksi esineiden internetissä olisi tarpeellista luoda kansainvälinen lainsäädännöllinen viitekehys ottaen huomioon IoT:n teknologiset erityispiirteet.

Esineiden internetin laitteet toimivat langattomasti sulautettuina ja läsnä olevina kokonaisuuksina ympäristössä, joten laitteisiin kohdistuvat hyökkäykset saattavat tapahtua fyysisesti tai salakuuntelua käyttämällä (Atzori ym., 2010). RFID-tageihin on vaikea toteuttaa tehokkaita tietoturvaratkaisuja ja varsinkin jos RFID-teknologioita käyttävä EPC-arkkitehtuuri muodostuu globaalisti palvelumalliksi, voi RFID-tagien haavoittuvuudet luoda merkittäviä liiketoiminnallisia, poliittisia ja yksityisyyteen liittyviä ongelmia ympäri maailman (Fabian & Günther, 2009). Erityisesti RFID-tunnisteiden ja mobiililaitteiden IoT-sovellusten mahdollistama paikkatietojen kerääminen ja analysointi voi heikentää käyttäjien yksityisyyden suojaa, jos paikkatiedot joutuvat ulkopuolisten toimijoiden käsiin (Aggarwal ym., 2013). RFID-tageja ei voida suojata tehokkaiden salausalgoritmien tai salasanojen avulla tagien vähäisen tehokkuuden vuoksi (Atzori ym., 2010). Tällöin tulisikin kehittää proaktiiviset menetelmät DDoS-hyökkäyksiltä ja haitallisilta ohjelmistoilta suojautumiseen sekä rakentaa luottamus sisäisesti, jotta laitteet voisivat itsenäisesti ja turvallisesti käsitellä luottamuksellisia ja kriittisiä tietoja esineiden internetin välityksellä (Bandyopadhyay & Sen, 2011). Vaikka IoT:n tiedonsiirto perustuu osittain jo monipuolisesti käytössä olleisiin protokolleihin ja standardeihin, luo esineiden internetin tehokkuusvaatimukset ja tiedonsiirron erityispiirteet laaja-alaisia tietoturvaongelmia, jos verkko ruuhkautuu tai viestien salaus ja osoitteen muodostus eivät toimi tarpeen mukaisesti (Granjal ym., 2015).

Tietoturva vaikuttaa IoT-järjestelmässä kokonaisvaltaisesti ja oleellista olisi kehittää luotettavia alustoja, vähentää kompleksisuutta sekä hallita tehokkaasti identiteettejä, yksityisyyttä ja erilaisten toimijoiden tunnistusta (Miorandi ym., 2012). Li ym. (2015) toteavat esineiden internetin laajempaa käyttöönottoa ja hyväksyntää kokonaisuudessaan kansainvälisesti rajoittavan merkittävästi kysymykset yksityisyyden, luotettavuuden ja tietoturvan näkökulmista.

### 3.1.3 Standardit

Standardointi on esineiden internetin kehityksen kannalta oleellista, jotta laitteet toimisivat yhdessä tehokkaasti, mutta esineiden internetin jatkuva laaja-alainen kehittyminen luo standardoinnista vaikean prosessin (Li ym., 2015). IoT-ympäristön sovellusrajapintoja, heterogeenisiä teknologiaratkaisuja ja laitteita sekä tietoturvanhallintaa varten täytyy toteuttaa yhdenmukaiset standardit, jotta esineiden internet voisi toimia monipuolisesti (Bandyopadhyay & Sen, 2011). Esineiden internet kytkee yhteen tai useampaan verkkoon lukuisia erilaista dataa tuottavia ja käyttäviä laitteita, mikä voi onnistuessaan kehittää liiketoimintaa monipuolisesti, mutta tekee kokonaisuuden hallinnasta haasteellista. Useiden teknologisten sovellusten hyödyntäminen luo IoT-ympäristöstä monimutkaisen ja sirpaleisen kokonaisuuden, mikä vaatii teknologisia standardien määrittelyn arkkitehtuurin eri kerroksille, varsinkin yhdenmukaisten väliohjelmistojen ja ohjelmistorajapintojen toteuttamiseksi (Li ym., 2015).

Eräänä olennaisena IoT-paradigman tavoitteena on lisätä automaatioominaisuuksia sekä kommunikointia erilaisten laitteiden välillä, mikä korostaa selkeiden ja yhtenäisten merkintätapojen ja sisältöjen käyttöä IoT-sovelluksissa (Whitmore ym., 2015). Koneiden väliselle kommunikoinnille ja tiedonsiirrolle (M2M) onkin keskeistä – standardien avulla mahdollistettu – datakeskeisyys, joka tehostaa datan reaaliaikaista ja hajautettua analysointia sekä skaalautuvuutta (Aggarwal ym., 2013). Lisäksi erityisille sovelluksille ominaisia toimialakohtaisia avoimia standardeja täytyisi asettaa, jotta esineiden internet voisi kehittyä laaja-alaisesti erilaisiin älyllisiin ratkaisuihin (Zanella ym., 2014).

## 3.2 Liiketoiminnalliset ja yleiset haasteet

Organisaation liiketoiminnan kannalta esineiden internetin hyödyntämisen haasteet liittyvät organisaation kustannuksiin ja resursseihin, kehittämissuunnitelman hallintaan, lainsäädäntöön ja IoT-sovellusten käyttämiseen. IoT-järjestelmän käyttöönoton kustannukset laitteiden hankkimisessa voi rajoittaa organisaation liiketoimintaa, sillä kokonaisvaltaisen IoT-ympäristöön sopivan protokollapinin toteuttaminen laitteisiin vaatii resursseja (Sheng ym., 2013). Avoimen esineiden internetin kokonaisuuden luomiseksi organisaation pitää toimia yhteistyössä useiden toimijoiden ja sidosryhmien kanssa, mikä voi rajoittaa IoT-järjestelmän laaja-alaista hallintaa organisaation näkökulmasta (Miorandi ym., 2012). Koska IoT-järjestelmä perustuu tehokkaaseen tiedonhallintaan, täytyy eri sektoreiden organisaatioiden pyrkiä avoimeen ja joustavaan tietojen jakamiseen, jotta IoT-palvelut olisivat toimintakykyisiä (Li ym., 2012). Esineiden internetin uudenlaisten liiketoimintamallien, kuten erilaisten pilvipalveluiden ja -alustojen, jatkuva kehittyminen voi aiheuttaa organisaatiolle haasteita uudistumisen ja tehokkaiden palveluiden muodostamisessa ja hallinnassa (Whitmore ym., 2015). Lisäksi sekä paikallisen että kansainvälisen lainsäädännön muuttuminen voi vaikuttaa yrityksen liiketoimintamahdollisuuksiin esineiden internetissä (Weber, 2010).

Kranz, Holleis & Schmidt (2010) toteavat, että olisi käytännöllistä luoda monipuoliset kehitystyökalut ja ohjelmointirajapinnat useita erilaisia IoT-toteutuksia, kehitysprosesseja ja sekä yksinkertaisia että monimutkaisia sovelluksia ja niiden kehittämistä varten. Innovaatioiden aikaansaamiseksi kaikenlaisia esteitä ja rajoitteita, kuten markkinoille pääsyä, tulisi vähentää sekä yhdistää erilaisia IoT-kehittämisstrategioita kustannustehokkaiden ja monipuolisten palveluiden mahdollistamiseksi (Li ym., 2015). IoT-laitteiden kehittämisessä tulisikin huomioida sulautetun tietotekniikan vaikutukset käyttämiseen, laitteiden luoma lisäarvo ja monipuolisuus, käyttäjien toiminnan eri ulottuvuudet sekä käyttäjiltä kerätyn datan laatu kontekstista riippuen (Kranz ym., 2010). Tulevaisuutta varten Whitmore ym. (2015) toteavat IoT:n käyttöönoton julkisella sektorilla, big datan integroinnissa, liiketoimintamallien luonnissa ja erityisissä konteksteissa tarvitsevan kattavammin tutkimustietoa, jotta esineiden internet voitaisiin suunnitella ja toteuttaa ratkaisemalla siihen liittyvät ongelmat järjestelmällisesti.



## 4 IOT-TIEDONHALLINNAN SOVELLUKSIA JA PALVELUITA

Esineiden internetin tiedonhallinnalle keskeistä on IoT-järjestelmän laitteiden hajautettu sijainti, heterogeeniset laiteverkostot, langaton tiedonsiirto sekä massiivisten tietomäärien käsittely (Atzori ym., 2010). Esineiden internetin sovellukset, kuten älykaupungit, luovat haasteellisia vaatimuksia datan ja tiedonsiirron laadulle, oikea-aikaisuudelle, tehokkuudelle sekä viestinnän perusteella tehtävälle analysoinnille ja päätöksenteolle (Rathore ym., 2016). IoT-tiedonhallinnan erilaisten sovellusten ja palveluiden täytyy tukea tiedon nopeaa, laadukasta ja monipuolista käsittelyä, jotta esineiden internetin toteutusten hyödyntäminen olisi kustannustehokasta ja kestävää pitkällä aikavälillä. Díaz ym. (2016) määrittelevät tiedonhallinnan keskeiseksi ratkaisuksi pilvidatakeskuksen, johon sisältyisi erilaisia alustoja liittyen tiedon tallentamiseen, eräajojen suorittamiseen, johtamiseen sekä oikea-aikaiseen prosessointiin. Suurten tietomassojen käsittelylle on keskeistä ohjelmistojen ja tietokantojen integrointi. Monipuoliset alustat, kuten avoimen lähdekoodin Apachen Hadoop, Spark ja Storm, ovatkin merkittäviä kokonaisuuksia, jotta esimerkiksi Apachen Cassandraan, HBaseen tai Amazonin S3-pilvipalveluun perustuvan tietovaraston hyödyntäminen olisi mahdollista IoT-järjestelmässä. (Díaz ym., 2016.) Seuraavaksi arvioidaan esineiden internetin tiedonhallinnan pilvipalveluita ja sovelluksia, jotka tukevat teknologioiden integrointia sekä mobiililaitteiden ja pilvilaskennan hyödyntämistä.

### 4.1 Pilvipalvelut

Gubbi ym. (2013) toteavat kaikkialla läsnä olevan Web 3.0:n kasvattavan merkittävästi datan kysyntää ja tarvetta internetin palveluiden luomisessa, mikä korostaa IoT-järjestelmän pilvikeskeistä (engl. cloud-centric) lähestymistapaa. Pilvipalveluna toimiva Aneka-alusta yhdistäisi yksityisen ja julkisen pilven, kuten Microsoft Azuren ja Amazonin EC2:n, tallennus- ja laskentaominaisuuksia, jotta esineiden internet ja tulevaisuuden Web 3.0 voisivat toimia tehokkaasti.

Aneka tukee tehtäväperusteista (engl. task programming), säikeisiin jakautuvaa (engl. thread programming) sekä MapReduce-ohjelmointia suurten datamäärien analysoinnissa. Lisäksi Aneka mahdollistaa palveluiden yksilöllisen muokkaamisen, valvonnan, automaattisen skaalautuvuuden sekä laskutuksen ja resurssien varaamisen. Erilaisten palveluiden tuottamiseen vaadittava laatu-taso voidaan varmistaa, sillä Anekan hallintaominaisuudet optimoivat resurssien jakamista ja aikataulutusta IoT-sovellusten käytön perusteella. Aneka-isäntäpalvelin ohjaa useita Aneka-instansseja ja -verkoston solmuja internetin välityksellä pilvipalvelun sisäisesti, jolloin palvelupyyntö toteutetaan kokoaamalla data useista eri lähteistä sekä hallinnoimalla Aneka-järjestelmää kokonaisuudessaan. (Gubbi ym., 2013.) Haasteiksi esimerkiksi Aneka-Azure hybridi-pilvessä muodostuukin useiden sovellusten tehokas ja tietoturvallinen hallinta. Erilaiset palvelunestohyökkäykset, datan virheellisyys, liiallinen heterogeenisuus tai puutteellisuus sekä tietoverkkojen tehoton yhteistoiminta voivat hidastaa palvelun toimintaa kokonaisvaltaisesti.

Älykkäiden mobiililaitteiden määrän kasvusta älylaitteita voidaan käyttää yhä monipuolisemmin IoT-järjestelmien kehittämisessä. Borgia, Bruno, Conti, Mascitti ja Passarella (2016) toteavat 2020-luvulla erilaisten pilvilaskennan alustojen ja informaatiokeskeisen verkottumisen (engl. information-centric networking, ICN) edistävät datan tavoitteellista ja tehokasta hyödyntämistä. Tiedonsiirron nopeus, kapasiteetti, tietoturva ja globaalius korostavat ICN- ja IoT-järjestelmien haasteita. Siirtämällä laskennan, tallennuksen ja älykkyyden ominaisuudet mobiiliverkkojen reunoilla oleville mobiililaitteille ja käyttäjille, voidaan järjestelmän kokonaistoimintakykyä parantaa merkittävästi. (Borgia ym., 2016.) Liu ym. (2017) jaottelevat mobiilireunapilven (engl. mobile edge cloud, MEC) järjestelmän käyttäjien älylaitteisiin, mobiilireunapilven palvelimiin, palveluiden jakamiseen ja palvelutietoiseen reitittämiseen (engl. service-aware routing) sekä internetin välityksellä toimiviin pilvi- ja datakeskuksiin. Borgia ym. (2016) jakavat MEC-alustan globaaliin ICN- ja pilvialustaan tiedonhallinnan ja palveluiden osittamisen perustaksi, mobiilipilviin tiedonsiirron, palveluiden jaottelun ja dynaamisen verkottumisen tukemiseksi sekä IoT-järjestelmän käyttäjiin ja laitteisiin, jotka luovan heterogeenistä raakadataa. Liu ym. (2017) mukaan mobiilireunapilvet eroavat tavallisista pilvijärjestelmistä siirtämällä tietojenkäsittelyn ominaisuuksia lähemmäksi käyttäjiä ja reagoimalla järjestelmässä tapahtuviin muutoksiin joustavammin, mikä edistää palvelujen saatavuutta ja vähentää viivettä tietojenkäsittelyssä. Toisaalta järjestelmä koostuu pienistä mikrodatakeskuksista sekä erilaisista ja hajautetuista laitteista, mikä luo haasteita kattavan MEC-arkkitehtuurin muodostamisessa, avoimen ja tehokkaan alustan kehittämisessä sekä protokollien ja algoritmien soveltamisessa. MEC-kokonaisuuden heterogeenisuus korostaa myös erilaisten resurssien ja tietojenkäsittelyn monipuolista hallintaa sekä muiden pilvipalveluiden ja mobiilijärjestelmien integrointia laaja-alaisesti. (Liu ym., 2017.)

Esineiden internetin ja ICN-järjestelmien toimintakyvyn ja tiedon monipuolisen hyödyntämisen tehostamiseksi Yue, Guo, Li, Asaeda ja Fang (2014) määrittelevät DataClouds-ratkaisun, joka yhdistää loogisia ja fyysisiä yhteisöjä tiedonkeruun ja jakamisen sekä erilaisten palveluiden tuottamisen mahdollistamiseksi. DataClouds-järjestelmä jakautuu käyttäjien loogiseen ja fyysiseen

tilaan. Käyttäjät ja käyttäjäryhmät jakavat samankaltaisia tarpeita ja ominaisuuksia loogisissa yhteisöissä, jotka yhdistetään fyysisesti tiedonsiirron väli­muistin, yhteisöjen kohtaamispisteiden sekä viestien edelleenvälittämisen lähet­täjä ja vastaanottajasolmuihin. DataClouds perustuu ICN- ja IoT-järjestelmien datakeskeisten palveluiden mahdollistamiseen, sillä ratkaisumalli jaottelee käyttäjät yhteisöihin heidän tietotarpeidensa ja kiinnostuksen kohteidensa mu­kaisesti, minkä perusteella DataClouds-tietoverkkokokonaisuus muodostetaan pilvipalveluihin. DataClouds-alusta rakentuu erilaisten yhteisöjen väliseen vuorovaikutukseen, jossa tietoturva ja eri organisaatioiden välinen yhteistyö vaikuttaa koko järjestelmän toimintaan. (Yue ym., 2014.) Alustan tiedot voi­daankin salata käyttämällä tehokkaita kryptaustekniikoita, mutta tiedonsiirron turvaaminen vaatii merkittäviä toimenpiteitä eri sidosryhmien osalta.

Perez ja Carrera (2015) määrittelevät servIoTicy-pilvipalvelualustan IoT-palveluiden tiedonhallinnan tueksi, mikä rakentuu useiden käyttäjien tietotarpeiden prosessointiin, data-analytiikkaan ja monipuolisiin kyselyihin, datakeskeisiin teknologiaratkaisuihin sekä REST-arkkitehtuurimalliin ohjelmointiraja-pinnassa. ServIoTicy-alustan toiminta perustuu abstraktien web- ja palveluobjektien, dataprosessointiputkien sekä erilaisten tietotarpeiden tilausten yhteistoimintaan. Pilvialustan arkkitehtuuri koostuu verkkokerroksen Jetty- ja Jackson-sovelluksista, datavirtojen hallinnan Apache STORM-alustasta, tietojen va­rastoinnin dokumenttipohjaisesta, NoSQL, CouchBase-tietokannasta, Elasticsearch-haku- ja indeksointikoneesta sekä REST-ohjelmointirajapinnan useista erilaisista integraatiomahdollisuuksista IoT-laitteiden välillä. ServIoTicyn to­teuttaminen vaatii tehokasta integrointia ja kehittämistä useiden sovellusten joustavan yhteistoiminnan mahdollistamiseksi, sillä yksittäisen teknologiarat­kaisun heikko toiminta voi rajoittaa koko järjestelmän palvelukykyä. (Perez & Carrera, 2015.)

IoT-tiedonhallinnan erilaiset pilvialustat korostavat esineiden internetin vaatimuksia suurten ja heterogeenisten tietomäärien käsittelyssä. Big data-analytiikan, äly- ja IoT-laitteiden, teknologisten sovellusten ja tietoverkkoratkai­sujen kehittymisen ansiosta esineiden internetin tiedonhallinnan pilvialustat muuttuvat jatkuvasti vastaamaan paremmin IoT-ympäristöjen tarpeita. Pilvi­ratkaisujen kehittämisessä onkin keskeistä määritellä laitteiden ja sovellusten integrointi, kustannustehokkuus ja palveluiden luonti sekä alustan joustavuus, tehokkuus ja skaalautuvuus.

## 4.2 Datakeskukset

Esineiden internetin monipuolisten sovellusten toiminta perustuu tehokkaaseen ja usein hajautettuun tietojenkäsittelyyn, mikä vaatii laiteverkostojen tuottaman datan tallennusta, analysointia ja hyödyntämistä keskitetysti. IoT-järjestelmien suorituskykyyn liittyy keskeisesti big data ja pilvilaskenta, sillä älykkäiden IoT-sovellusten ja -palveluiden mahdollistaminen rakentuu suurten tietomäärien ja erilaisten pilvialustojen joustavaan ja integroituvaan yhteistoimintaan (Díaz ym., 2016). Atzori ym. (2010) määrittelevät IoT-järjestelmän muodostuvan esi-

neiden näkökulmasta heterogeenisistä matalatehoisista ja älyllisistä laitteista, joiden tietojenkäsittelyn, tallentamisen ja analysoinnin ominaisuudet voivat olla hyvin erilaisia. Nykyään datakeskukset sekä niihin toteutetut monipuoliset teknologiset kyvykkyydet ovatkin Barin ym. (2013) mukaan olleet keskeisenä elementtinä suurten tietomäärien käsittelyssä ja internet-sovellusten kehittämisessä. Toisaalta Al-Fares, Loukissas ja Vahdat (2008) toteavat datakeskusten toimintakyvyn heikkenevän kokonaisvaltaisesti, mikäli datakeskusarkkitehtuurin kaikki osa-alueet eivät pysty palvelemaan tasapuolisesti, joustavasti ja skaalautuvasti. Hua, Liu ja Jiang (2014) korostavat pilvidatakeskusten merkittävyyttä, sillä nykyään datakeskukset toteuttavat myös sähköposti-, tiedonhaku- ja web-palveluita, missä haasteeksi muodostuu tiedonsiirron, skaalautuvuuden sekä datan ja datakeskusten välinen sijoittuneisuus verkossa. IoT-tiedonhallinnan kannalta erilaiset datakeskusratkaisut ovatkin keskeisessä roolissa, jotta IoT-palveluille tärkeitä big data-analytiikkaan ja tiedonlouhintaan liittyviä toteutuksia voitaisiin käyttää hyödyllisesti (Akhbar ym., 2016).

Hua ym. (2014) määrittelevät semanttisesti toimivan ANTELOPE-datakuution pilvidatakeskuksia varten. ANTELOPE-datakuutio rakentuu abstraktisti suorakulmaisista laskentayksiköistä – särmiöistä, jotka on liitetty järjestelmällisesti yhteen tiettyjen laskennallisia ominaisuuksia vertailemalla. Datakuutiot voidaan jaotella, huomioimalla datan merkitykset, esimerkiksi mittaus tulosten ajan, käytettyjen palvelinten sekä käyttöjärjestelmän kirjoitus- ja lukuoperaatioiden perusteella. Datakuutio erittelee kerätyn datan ominaisuudet tehokkaasti hyödyntämällä paikallista hajautusta (engl. locality sensitive hashing, LSH), mikä mahdollistaa datan mallintamisen dynaamisesti ja tasapainotetusti suorakulmioista muodostuvaan R-puuhun (engl. rectangle tree, R-tree). Jotta ANTELOPE-datakuutio voisi välttää hitaita lineaarisia hakuja toimissaan, on data indeksointi ja ryhmittely toteutettu tilastollisia menetelmiä ja raja-arvoja hyödyntämällä. Vertailtaessa ANTELOPE:n toimintakykyä yleisesti käytössä olevien datakeskusmallien kanssa, Hua ym. (2014) toteavat ANTELOPE-datakuution palvelevan tehokkaammin, kun datakeskusten pitää ymmärtää datan erilaisia merkityksiä verkkoarkkitehtuurin optimoinnin kannalta. ANTELOPE on datakeskeisyyteen ja semanttisuuteen pohjautuva malli – etenkin IoT- ja big data-järjestelmiä varten, sillä ANTELOPE edistää datakeskusten datan sijainnin ja rakenteen tehokasta ylläpitoa, mahdollistaa erilaisten tiedonsiirtotapojen käytön sekä tukee kokonaisvaltaista skaalautuvuutta. Toisaalta kustannusten vähentämiseksi, datakeskuksen muodostamisessa, ANTELOPE hyödyntää semanttisesti osittaista materialisaatiota (engl. partial materialization), joka voi heikentää kyselyjen tarkkuutta. (Hua ym., 2014.)

Esineiden internetin tiedonhallinta voidaan toteuttaa myös joustavasti mobiililaiteverkostojen sisäisesti. Borgia ym. (2016) korostavat informaatiokeskeisen verkottumisen merkitystä sekä käyttäjien äylaitteiden laskennallisten ominaisuuksien hyödyntämistä IoT-tiedonhallinnan, etenkin mobiilireunapilvien alustaratkaisuissa. Akhbar ym. (2016) mukaan laajojen datakeskuskokonaisuuksien ohella olisi hyödyllistä kohdentaa tietojenkäsittelyn laskentaominaisuuksia useille pienemmille ja lähempänä toisiaan oleville laiteyksiköille eli nanodatakeskuksille (NaDa). Valancius, Laoutaris, Massoulié, Diot ja Rodriguez (2009) toteavat nanodatakeskuksissa palveluiden ja sisältöjen tallentamisen

tiedonsiirron yhdysväyliin, datakeskusten sijaan, edistävän erityisesti tiedonhallinnan kustannustehokkuutta ja vähentävän energiankulutusta. NaDa-arkkitehtuuri rakentuu hajautettuun palvelualustaan, jossa tietojen varastointi ja käsittely toteutetaan verkon reunoilla olevissa pienimuotoisissa palvelimissa. Tällöin nanodatakeskusratkaisun palvelimista ja tiedonsiirrosta vastaa internet-palveluntarjoaja, sillä NaDa-kokonaisuus perustuu vertaisverkkoihin (engl. peer-to-peer, P2P) ja niistä muodostuviin vertaisverkkojärjestelmiin. Nanodatakeskukset voisivat edistää IoT-ympäristön tiedonhallinnan itsestään skaalautuvuutta, energiatehokkuutta sekä palvelunluonnin läheisyyttä verkossa. Toisaalta palveluntarjoajien ja asiakkaiden välisten sopimusten tekeminen NaDa-järjestelmän reitittimien, energiankulutuksen ja sovellusten käytön osalta on keskeistä. Tiedonsiirron kaistanleveys ja kapasiteetti ovat merkittävimpiä NaDa-kokonaisuuden resursseja, joiden toiminnan ylläpito ja kehittäminen vaikuttavat koko järjestelmän suorituskykyyn. (Valancius ym., 2009.) Nanodatakeskusten hyödyntämistä erilaisissa IoT-sovelluksissa täytyisikin testata ja simuloida laaja-alaisesti, jotta haasteet yksityisyydessä, tietoturvassa, palveluiden laadussa sekä tietojen tallennuksessa ja palauttamisessa voitaisiin ratkaista (Akhbar ym., 2016).

Esineiden internetin tiedonhallinta vaatii kattavaa ja tehokasta tietojen varastointia, joka tukisi erilaisten laitteiden välistä tiedonsiirtoa ja tiedonkeruuta sekä hyödyllisten palveluiden ja sovellusten luontia tietojenkäsittelyn ansiosta. Datakeskukset IoT-tiedonhallinnan mahdollistajina ovat keskeisessä osassa, sillä esineiden internetin erityispiirteet asettavat merkittäviä vaatimuksia ja tarpeita tietojen tallennuksen toteuttamiselle. Esineiden internetissä datakeskukset eivät ole ainoastaan suuria tietovarastoja vaan myös data-analytiikkaan ja palveluiden luontiin liittyviä skaalautuvia kokonaisuuksia. Uudenlaiset mallit datakeskusten suorituskyvyn edistämiseksi voivat edesauttaa IoT-tiedonhallinnan monipuolista toteuttamista, mutta esineiden internetin globaalien kehittymisen vaikutuksesta tietojen tallennuksen ja käsittelyn toteutuksien täytyy kehittyä yhtäläisesti.

Kokonaisuudessaan esineiden internetin tiedonhallinnan kannalta merkittävimpiä pilvialusta- ja datakeskusratkaisuja on haasteellista arvioida tutkimuskirjallisuuden perusteella. Yleisesti tutkimusartikkeleissa käsitellyt sovellukset pystyvät toimimaan ja palvelemaan tehokkaasti erilaisia asiakkaita, tietomääriä ja laitekokonaisuuksia skaalautuvasti tutkimusympäristössä – usein teoreettisesti. IoT-järjestelmään sopivia sovelluksia ja palveluita kehitetään jatkuvasti moniulotteisten ongelmien ratkaisemiseen, mutta ratkaisujen kokonaisvaltaista hyödyllisyyttä, IoT-ympäristön toimintakyvyn edistämisessä, on mahdollista arvioida vasta käytännön toteutusten ja reaali maailman testauksen ansiosta. Esineiden internetin tiedonhallinnan sovellukset ja palvelut perustuvat monipuolisesti erilaisten hyötyjen saavuttamiseen, mutta toteutuksiin liittyy keskeisesti myös merkittäviä haasteita, jotka täytyy analysoida ja ratkaista järjestelmien kehittämisessä (TAULUKKO 2).

TAULUKKO 2 IoT-tiedonhallinnan sovelluksia ja palveluita

Sovellus tai palvelu	Tyyppi	Hyödyt	Haasteet	Lähteet
Aneka	Pilvialusta	Yksityisen ja julkisen pilven yhdistäminen, useat ohjelmointimallit, palveluiden monipuoliset hallintatyökalut	Erilaisten sovellusten integrointi, tietoturva, palvelun saatavuus	Gubbi ym., 2013
Mobiilireunapilvi	Pilvialusta	Palvelujen läheisyys, saatavuus ja muokattavuus; informaatiokeskeinen verkottuminen	Mikrodatakeskusten ja alustan hallinta; protokollien ja algoritmien luominen	Borgia ym., 2016; Liu ym., 2017
DataClouds	Pilvialusta	Loogisten ja fyysisten yhteisöjen verkottuminen; datakeskeisten palveluiden luominen	Tiedonsiirron ja sisältöjen tietoturva; käyttäjien välinen yhteistyö	Yue ym., 2014
servIoTicy	Pilvialusta	Monipuoliset teknologiaratkaisut, data-analytiikka, tietotarpeisiin vastaaminen	Useiden sovellusten integrointi ja kehittäminen kokonaisuuden kannalta	Perez & Carrera, 2015
ANTELOPE	Datakeskus	Semanttisesti toimiva datakuutio, verkkoarkkitehtuurin tehokas optimointi, skaalautuvuus	Laskennallisten yksiköiden ylläpito, kustannukset, kyse-lyjen täsmällisyys	Hua ym., 2014
Nanodatakeskus	Datakeskus	Tiedonhallinnan kustannustehokkuus ja läheisyys, energiankulutuksen optimointi, skaalautuvuus	Palveluntarjoajien ja käyttäjien välinen yhteistyö; tietoturva, palveluiden laatu, tiedonsiirron ylläpito, yksityisyys	Valancius ym., 2009; Akhbar ym., 2016

## 5 YHTEENVETO

Tutkielmassa havainnollistettiin kirjallisuuskatsauksen avulla tiedonhallintaan liittyviä kokonaisuuksia, erityispiirteitä ja haasteita esineiden internetin aikakaudella. Lisäksi tutkielmassa käsiteltiin esineiden internetin tiedonhallinnan mahdollistavia pilvialusta- ja datakeskusratkaisuja. Vaikka esineiden internet on suhteellisen tuore käsite, teknologioita kokoava paradigma ja tutkimuskohde, sisältää esineiden internetin tutkimuskirjallisuus laaja-alaisesti erilaisia ulottuvuuksia IoT-ympäristön ymmärtämiseksi. Lähdeaineistosta valittiin laadullisesti merkittävimmät, mielenkiintoisimmat ja tulevaisuuden kehityksen kannalta oleelliset kokonaisuudet yksityiskohtaisemman tarkastelun kohteeksi.

Kandidaatintutkielma jakautuu johdannon lisäksi kolmeen sisältöluokkaan ja yhteenvetoon. Ensimmäisessä sisältöluvussa käsiteltiin esineiden internetin tiedonhallintaa DMBOK2-viitekehyksen (Cupoli ym., 2014) mukaisesti. Ensimmäiseen tutkimusongelmaan vastattiin määrittelemällä esineiden internetin tiedonhallinnan ominaispiirteiden liittyvän keskeisesti kehitettävien ratkaisujen, tietorakenteiden, tietoturvan, tiedon integroinnin ja erilaisten protokollien hallintaan. Tiedonhallinnan viitekehys tuki tulosten jaottelua tehokkaasti, mutta tietyt osa-alueet, kuten dokumentointi ja sisältö sekä metatieto, jäivät vähemmälle painoarvolle tutkielmassa, johtuen lähdeaineiston puutteellisuudesta, mikä on tutkielman keskeisin rajoite. Toisaalta IoT-järjestelmien toimivuudesta ei yleisesti löytynyt merkittävää lähdeaineistoa, mikä voi johtua IoT-toteutusten minimaalisuudesta sekä tutkimusyhteistyön haasteellisuudesta IoT-yritysten kanssa. Lähdeaineisto ensimmäisessä sisältöluvussa tuki monipuolisesti eri näkökulmien käsittelyä, mutta tutkimusartikkelit vastasivat tavallisesti hyvin suppeasti ja melko abstraktisti tiedonhallinnan eri osa-alueisiin. Esineiden internetin tutkimus on keskittynyt ensisijaisesti arkkitehtuurien, tietoturvan, integroinnin ja tietomäärien hallinnan ongelmien ratkaisuun, joten olisikin merkittävää tutkia kattavammin IoT-tiedonhallintaa kokonaisuudessaan sekä muista keskeisistä näkökulmista.

Toisessa sisältöluvussa vastattiin toiseen tutkimusongelmaan selvittämällä perusteellisesti teknologisten toteutusten, tietoliikenteen, tietoturvan ja yksityisyyden sekä käytettävien standardien asettavan erilaatuisia haasteita esineiden internetin tiedonhallinnalle. Teknologisen ja yhteiskunnallisen kehityksen vai-

kutuksesta haasteiden määrittely on ongelmallista, sillä IoT-tutkimusten hyödyllisyys ja relevanttius tietyssä kontekstissa voi muuttua hyvinkin lyhyellä aikavälillä. Yleisesti tässä kandidaatintutkielmassa keskityttiin keskeisimpien haasteiden käsittelyyn, yksilöllisten – teknologisten tai liiketoiminnallisten – ongelmien sijasta, sillä tutkielma painotti IoT-tiedonhallinnan merkittävimpien kokonaisuuksien hyötyjä ja kehitystarpeita laaja-alaisesti. Lähdeaineisto sisälsi kattavasti IoT-ympäristön haasteiden arviointia, mutta haasteiden käsittely tutkimusartikkeleissa keskittyi erilaisten haasteiden erittelyyn ja vertailuun, mikä osaltaan vaikutti rajoitteena. IoT-tiedonhallinnan kehittymisen kannalta olisikin oleellista käsitellä mahdollisia ratkaisumalleja IoT-järjestelmän ongelmille sekä muodostaa selkeästi tarpeita ja vaatimuksia tutkimukselle tulevaisuudessa.

Lopuksi kolmannessa sisältöluvussa määriteltiin esineiden internetin tiedonhallinnan erityispiirteiden ja haasteiden perusteella IoT-tiedonhallintaan soveltuvien pilvialusta- ja datakeskusratkaisujen ominaisuudet, hyödyt ja rajoitteet IoT-järjestelmän näkökulmasta. Tutkimuskirjallisuus sisältää useita ratkaisuja liittyen pilvipalveluihin ja datakeskuksiin, mutta tähän tutkielmaan valittiin viimeisimmät ja toisistaan poikkeavimmat kokonaisuudet, jotka liittyivät ominaisuuksiltaan muihin tutkielman tutkimuskysymyksiin keskeisesti. Lähdeaineistossa ei arvioitu selkeästi IoT-tiedonhallinnan ratkaisujen käytännöllisyyttä ja merkittävimpiä haasteita, ja tutkielmassa käsitellyt sovellukset olivatkin mahdollisia ratkaisumalleja ja viitekehyksiä tiedonhallinnan tukemiseksi. Lähdeaineiston perusteella on vaikea arvioida, kuinka tehokkaasti ratkaisut ovat toteutettavissa kustannuksien, laitteistojen ja järjestelmien integroinnin sekä sopimusten hallinnan osalta. IoT-tiedonhallinnan ratkaisujen arviointi, vertailu ja kehittäminen vaatii tulevaisuudessa monipuolista yhteistyötä ja reaaliaikailmaan perustuvan datan hyödyntämistä.

Kandidaatintutkielman tiedonhankinta toteutettiin käyttämällä Scopus-tietokantaa ja Google Scholar-hakukonetta. Hakulausekkeet muodostettiin pääasiallisesti tutkielman tiivistelmässä mainittujen avainsanojen sekä tutkielman sisältöluvuissa käsiteltyjen keskeisimpien aihekokonaisuuksien yhdistelmillä. Lähdeaineiston tieteellinen relevanttius, arvo ja hyödyllisyys arvioitiin käyttämällä Julkaisufoorumi-palvelun Julkaisukanavahakua, jossa tieteellisten julkaisujen laadullinen taso on määritelty kattavasti. Lähdeaineiston valinnassa painotettiin tieteellisten julkaisujen laatutasoa, julkaisuvuotta, julkaisuun tehtyjen sitaattien määrää sekä tulosten hyödyllisyyttä ja selkeyttä tutkimusongelmien ratkaisemisen näkökulmasta.

Esineiden internetin tiedonhallinnan erityispiirteet kehittyvät dynaamisesti ja moniulotteisesti, joten esineiden internetiin liittyvien tutkimusten tarve kasvaa jatkuvasti. Hyödylliset jatkotutkimusaiheet IoT-tiedonhallinnan kannalta voisivat keskittyä yksityiskohtaisemmin sovellusten, järjestelmien tai yksittäisen tiedonhallinnan osa-alueen analysointiin. Tutkielman lähdeaineiston perusteella esineiden internetin tutkimuksen trendit ovat keskittyneet enimmäkseen haasteiden, mallien ja mahdollisuuksien arviointiin. IoT-järjestelmän käytännön toteutusten ja vähäisemmin tutkittujen tiedonhallinnon osa-alueiden, kuten metadatan hallinnan ja dokumentoinnin, merkitystä olisikin hyödyllistä tutkia perusteellisesti. Käyttäjien näkökulmasta olisi merkittävää arvioida kokonaisvaltaisesti IoT-tiedonhallinnan erityispiirteiden vaikutuksia yksilön, yri-



tyksen, yhteiskunnan ja globaalien markkinoiden tasolla. Esineiden internet mahdollistaa erilaisten innovaatioiden kehittämisen, mutta IoT-järjestelmien potentiaalia ei ole vielä hyödynnetty käytännössä laajemmin. Tämän vuoksi olisi tarpeellista käsitellä IoT-tiedonhallinnan aikaansaamia muutoksia, vaatimuksia ja ongelmia moniulotteisesti sekä arvioida IoT-innovaatioiden käyttöönoton rajoitteita toimialojen ja toimintaympäristöjen ominaisuuksien kannalta. Lisäksi olisi hyödyllistä määritellä, mitkä ominaisuudet rajoittavat esineiden internetin omaksumista ja millaisia haasteita täytyy ratkaista tulevaisuudessa IoT-järjestelmien toiminnan edistämiseksi pitkällä aikavälillä. Yleisesti jatkotutkimusaiheet voisivat pohjautua esineiden internetin, big datan ja datakeskeisyyden sovellusten käyttöönoton analysointiin eri toimialoilla. Lähdeaineiston perusteella useat IoT-tutkimukset ovat keskittyneet teoreettisten mallien ja sovellusten määrittelyyn, jolloin ratkaisujen suhdetta käytännön toteutuksiin ei ole käsitelty perusteellisesti. Käytännölliset tutkimukset voisivat painottua esimerkiksi RFID-tagien, tiedonhallinnan sovellusten, arkkitehtuurien tai tiedonsiirron teknologisten ratkaisujen vertailuun IoT-järjestelmistä kerätyn datan perusteella. Teknologisen toimintakyvyn arvioinnin ohella jatkotutkimukset voisivat käsitellä organisaatioiden toimintatapojen, arvonluonnin ja päätöksenteon muutoksia IoT-ratkaisujen, big datan ja datakeskeisyyden yhteisvaikutuksesta. Esineiden internet perustuu monipuolisten teknologioiden integrointiin, joka vaatii myös liiketoimintamallien, sidosryhmien välisen yhteistyön ja työtehtävien monipuolista kehittymistä. Esineiden internetin tutkimuksen pitäisi-kin huomioida kattavammin IoT-järjestelmien ominaispiirteitä ihmisten ja fyysisestä ympäristöstä kerättävän datan merkitysten näkökulmista.

Esineiden internet on laaja ja mielenkiintoinen paradigma, joka sisältää arkisia, monikäyttöisiä ja futuristisia sovelluksia niin yksilöiden, organisaatioiden kuin yhteiskuntien toiminnan edistämiseksi. IoT-järjestelmät tukevat älylaitteiden laaja-alaista ja verkostomaista yhteistoimintaa, mikä voi olla perustana tehokkaille, monikäyttöisille ja itsestään optimoituville datakeskeisille laiteverkostoille (Atzori ym., 2010). IoT-ympäristön luominen vaatiikin kokonaisvaltaista lähestymistapaa heterogeenisten laiteverkostojen vuorovaikutuksen kehittämiseksi. Internetin välityksellä toimivien älylaitteiden ja sensoreiden sekä niiden tuottaman datan määrä kasvaa jatkuvasti, mikä vaatii big data-analytiikkaa ja datakeskeistä lähestymistapaa IoT-järjestelmän toimintakyvyn ylläpitämiseksi (Aggarwal ym., 2013). Miorandi ym. (2012) määrittelevätkin IoT-tutkimuksen kannalta osa-alueiksi hajautettujen järjestelmien ja älykkyyden, tietojenkäsittelyn, tiedonvälityksen sekä identifiointin tutkimisen. Tämä kandidaatintutkielma keskittyi esineiden internetin palveluiden kannalta keskeisimmän resurssin – datan – laaja-alaisen hallinnan käsittelyyn. Menestyksekkäs tiedonhallinta mahdollistaa yritysten kilpailukykyä ja asiakkaiden tyytyväisyyttä edistävän arvonluonnin IoT-järjestelmissä, joissa palveluiden skaalautuvuus ja tehokkuus perustuvat aivan uudentyyliin vaatimukseen ja tarpeisiin. Esineiden internetin tietojenkäsittelyn siirtyessä lähemmäksi käyttäjiä ja fyysistä ympäristöä voidaan hyödyntää reaaliaikaisesti perustuvaa dataa palveluiden muodostamisessa yhä monipuolisemmin. Yksityisyyden ja tietoturvan ylläpitäminen ovat kuitenkin IoT-tiedonhallinnan merkittävimmät ominaispiirteet, joiden hallinta vaatii erityistä tehokkuutta ja huolellisuutta.

## LÄHTEET

- Ackoff, R. L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16(1), 3-9.
- Aggarwal, C. C., Ashish, N. & Sheth, A. (2013). The internet of things: A survey from the data-centric perspective. *Managing and Mining Sensor Data*, 383-428.
- Akhbar, F., Chang, V., Yao, Y. & Méndez Muñoz, V. (2016). Outlook on moving of computing services towards the data sources. *International Journal of Information Management*, 36(4), 645-652.
- Al-Fares, M., Loukissas, A. & Vahdat, A. (2008). A scalable, commodity data center network architecture. *In ACM SIGCOMM Computer Communication Review* (Vol. 38, No. 4, pp. 63-74). ACM.
- Alavi, M. & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107-136.
- Atzori, L., Iera, A. & Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, 54(15), 2787-2805.
- Bandyopadhyay, D. & Sen, J. (2011). Internet of things: Applications and challenges in technology and standardization. *Wireless Personal Communications*, 58(1), 49-69.
- Bari, M. F., Boutaba, R., Esteves, R., Granville, L. Z., Podlesny, M., Rabbani, M. G., Zhang, Q. & Zhani, M. F. (2013). Data center network virtualization: A survey. *IEEE Communications Surveys & Tutorials*, 15(2), 909-928.
- Bellavista, P., Corradi, A., Fanelli, M. & Foschini, L. (2012). A survey of context data distribution for mobile ubiquitous systems. *ACM Computing Surveys (CSUR)*, 44(4), 24.
- Bonomi, F., Milito, R., Zhu, J. & Addepalli, S. (2012). Fog computing and its role in the internet of things. *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, 13-16.
- Borgia, E., Bruno, R., Conti, M., Mascitti, D. & Passarella, A. (2016). Mobile edge clouds for information-centric IoT services. *2016 IEEE Symposium on Computers and Communication (ISCC)*, 422-428.
- Brewer, E. A. (2000). Towards robust distributed systems. *PODC* (Vol. 7).
- Chalasani, S. & Boppana, R. V. (2007). Data architectures for RFID transactions. *IEEE Transactions on Industrial Informatics*, 3(3), 246-257.
- Chen, H., Chiang, R. H. & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188.
- Chen, M., Mao, S. & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications*, 19(2), 171-209.
- Chianese, A., Marulli, F., Piccialli, F., Benedusi, P. & Jung, J. E. (2017). An associative engines based approach supporting collaborative analytics in the internet of cultural things. *Future Generation Computer Systems*, 66, 187-198.

- Cook, D. J. & Das, S. K. (2012). Pervasive computing at scale: Transforming the state of the art. *Pervasive and Mobile Computing*, 8(1), 22-35.
- Da Xu, L., He, W. & Li, S. (2014). Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4), 2233-2243.
- Díaz, M., Martín, C., & Rubio, B. (2016). State-of-the-art, challenges, and open issues in the integration of Internet of things and cloud computing. *Journal of Network and Computer Applications*, 67, 99-117.
- Dijkman, R. M., Sprenkels, B., Peeters, T. & Janssen, A. (2015). Business models for the Internet of Things. *International Journal of Information Management*, 35(6), 672-678.
- Ebert, C., Gallardo, G., Hernantes, J., & Serrano, N. (2016). DevOps. *IEEE Software*, 33(3), 94-100.
- Fabian, B. & Günther, O. (2009). Security challenges of the EPCglobal network. *Communications of the ACM*, 52(7), 121-125.
- Fazio, M. & Puliafito, A. (2015). Cloud4sens: A cloud-based architecture for sensor controlling and monitoring. *IEEE Communications Magazine*, 53(3), 41-47.
- Granjal, J., Monteiro, E. & Sa Silva, J. (2015). Security for the internet of things: A survey of existing protocols and open research issues. *IEEE Communications Surveys and Tutorials*, 17(3), 1294-1312.
- Grolinger, K., Higashino, W. A., Tiwari, A. & Capretz, M. A. (2013). Data management in cloud environments: NoSQL and NewSQL data stores. *Journal of Cloud Computing: Advances, Systems and Applications*, 2(1), 1.
- Gubbi, J., Buyya, R., Marusic, S. & Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645-1660.
- Hua, Y., Liu, X. & Jiang, H. (2014). ANTELOPE: A semantic-aware data cube scheme for cloud data center networks. *IEEE Transactions on Computers*, 63(9), 2146-2159.
- Khatri, V. & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148-152.
- Kokkonen, J. (2017). *Nanoscale sensor networks : the THz band as a communication channel*. Tieto- ja sähkötekniikan väitöskirja. Tieto- ja radioliikennetekniikan tutkimusyksikkö. Oulun yliopisto.
- Kranz, M., Holleis, P. & Schmidt, A. (2010). Embedded interaction: Interacting with the internet of things. *IEEE Internet Computing*, 14(2), 46-53.
- Leavitt, N. (2010). Will NoSQL databases live up to their promise? *Computer*, 43(2), 12-14.
- Lenzerini, M. (2002). Data integration: A theoretical perspective. *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, 233-246.
- Li, S., Da Xu, L. & Zhao, S. (2015). The internet of things: A survey. *Information Systems Frontiers*, 17(2), 243-259.
- Li, Y., Hou, M., Liu, H. & Liu, Y. (2012). Towards a theoretical framework of strategic decision, supporting capability and information sharing under the context of internet of things. *Information Technology and Management*, 13(4), 205-216.

- Liu, H., Bolic, M., Nayak, A. & Stojmenovic, I. (2008). Taxonomy and challenges of the integration of RFID and wireless sensor networks. *IEEE network*, 22(6), 26-35.
- Liu, H., Eldarrat, F., Alqahtani, H., Reznik, A., de Foy, X. & Zhang, Y. (2017). Mobile Edge Cloud System: Architectures, Challenges, and Approaches. *IEEE Systems Journal*.
- Mesiti, M. & Valtolina, S. (2014). Towards a user-friendly loading system for the analysis of big data in the internet of things. *Computer Software and Applications Conference Workshops (COMPSACW), 2014 IEEE 38th International* (pp. 312-317). IEEE.
- Miorandi, D., Sicari, S., De Pellegrini, F. & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7), 1497-1516.
- Mithas, S., Ramasubbu, N. & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly: Management Information Systems*, 35(1), 237-256.
- Perera, C., Zaslavsky, A., Christen, P. & Georgakopoulos, D. (2014). Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, 16(1), 414-454.
- Perez, J. L. & Carrera, D. (2015). Performance characterization of the servIoTicy API: An IoT-as-a-service data management platform. *Proceedings - 2015 IEEE 1st International Conference on Big Data Computing Service and Applications, BigDataService 2015*, 62-71.
- Qin, Y., Sheng, Q. Z., Falkner, N. J. G., Dustdar, S., Wang, H. & Vasilakos, A. V. (2016). When things matter: A survey on data-centric internet of things. *Journal of Network and Computer Applications*, 64, 137-153.
- Rathore, M. M., Ahmad, A., Paul, A. & Rho, S. (2016). Urban planning and building smart cities based on the internet of things using big data analytics. *Computer Networks*, 101, 63-80.
- Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180.
- Sheng, Z., Yang, S., Yu, Y., Vasilakos, A. V., McCann, J. A. & Leung, K. K. (2013). A survey on the ietf protocol suite for the internet of things: Standards, challenges, and opportunities. *IEEE Wireless Communications*, 20(6), 91-98.
- Sun, Y. & Jara, A. J. (2014). An extensible and active semantic model of information organizing for the internet of things. *Personal and Ubiquitous Computing*, 18(8), 1821-1833.
- Valancius, V., Laoutaris, N., Massoulié, L., Diot, C. & Rodriguez, P. (2009). Greening the internet with nano data centers. *In Proceedings of the 5th international conference on Emerging networking experiments and technologies* (pp. 37-48). ACM.
- Weber, R. H. (2010). Internet of Things–New security and privacy challenges. *Computer Law & Security Review*, 26(1), 23-30.
- Welbourne, E., Battle, L., Cole, G., Gould, K., Rector, K., Raymer, S., Balazinska, M. & Borriello, G. (2009). Building the internet of things using RFID: the RFID ecosystem experience. *IEEE Internet Computing*, 13(3), 48-55.

- Whitmore, A., Agarwal, A. & Da Xu, L. (2015). The internet of things - A survey of topics and trends. *Information Systems Frontiers*, 17(2), 261.
- Wong, A. K. Y., Ray, P., Parameswaran, N. & Strassner, J. (2005). Ontology mapping for the interoperability problem in network management. *IEEE Journal on selected areas in Communications*, 23(10), 2058-2068.
- Yue, H., Guo, L., Li, R., Asaeda, H. & Fang, Y. (2014). DataClouds: Enabling community-based data-centric services over the internet of things. *IEEE Internet of Things Journal*, 1(5), 472-482.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L. & Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1), 22-32.
- Zimmermann, A., Schmidt, R., Sandkuhl, K., Wißotzki, M., Jugel, D. & Möhring, M. (2015). Digital enterprise architecture-transformation for the internet of things. *2015 IEEE 19th International Enterprise Distributed Object Computing Workshop*.

## KAUPALLISET LÄHTEET

- Bluetooth (2016, 7. joulukuuta). Bluetooth 5 Now Available. Haettu 6.1.2017 osoitteesta  
<https://www.bluetooth.com/news/pressreleases/2016/12/07/bluetooth-5-now-available>
- Cupoli, P., Earley, S. & Henderson, D. (2014). DAMA-DMBOK2 Framework. Haettu 22.11.2016 osoitteesta  
<https://www.dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf>
- Greenough, J. (2016). How the 'Internet of Things' will impact consumers, businesses, and governments in 2016 and beyond. Haettu 27.3.2017 osoitteesta <http://nordic.businessinsider.com/how-the-internet-of-things-market-will-grow-2014-10?r=US&IR=T>